University of Alberta

FinO Structural and Functional Studies Reveal the Molecular Basis for F-like Plasmid Fertility Inhibition

by

Alexandru Florin Ghetu ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Biochemistry

Edmonton, Alberta

Spring 2002

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-68577-2

Canada

University of Alberta

Library Release Form

Name of Author: Alexandru Florin Ghetu

Title of Thesis: FinO Structural and Functional Studies Reveal the Molecular Basis for F-like Plasmid Fertility Inhibition

Degree: Doctor of Philosophy

Year this Degree Granted: 2002

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

JAN 30/2002

Alex Ghetu

151 Aurora Heights Drive
Aurora, Ontario
L4G 2X1, Canada

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **FinO Structural and Functional Studies Reveal the Molecular Basis for F-like Plasmid Fertility Inhibition** submitted by Alexandru Florin Ghetu in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

JNMark Glover
J. N. Mark Glover
(supervisor)

Laura Frost
Laura S. Frost

MJames
Michael N.G. James

Luis B. Agellon

Jack F. Greenblatt
(external)

JAN 30/2002
Date of Approval

Abstract

F-plasmid mediated conjugation requires the gene products of the plasmid encoded multicistronic transfer (*tra*) operon. Transcription of this operon is under the tight regulatory control of the transcriptional activator TraJ. Inhibition of conjugation requires the two components of the FinOP system: the protein FinO and the FinP RNA that is antisense to the 5' UTR of *traJ* mRNA. FinP contains two stem-loops, SLI and SLII, with the minimal binding target for FinO being SLII. FinO protects FinP from cellular RNases and promotes the duplex formation between FinP and *traJ* mRNA. Duplexing occludes a RBS on *traJ* mRNA, preventing expression of TraJ and thereby blocking conjugation.

Using limited proteolysis, EMSA and CD-spectroscopy, we could define two RNA binding regions within FinO; an N-terminal binding region located between residues 26 to 61 and a larger binding region consisting of a structurally stable core (residues 62-174) and a flanking C-terminal tail (residues 175-186).

The crystal structure of FinO was determined to 2.0 Å resolution. FinO is an elongated molecule resembling a fist with an extended index finger. One FinO-binding region is associated with a highly solvent exposed N-terminal helix (residues 35-67), whereas the other one is formed by a contiguous surface between a structural core and the tip of a long C-terminal helix packed against the base of the N-terminal helix. Both binding regions contain a basic patch that we suggest mediates RNA interactions.

Site-specific crosslinking and gelFRET assays were used to define the interactions that occur between FinO and SLII. Crosslinking between RNA and cysteines introduced at specific sites on FinO confirmed that two basic patches on the protein interact with

SLII, whereas the acidic region on the base of FinO does not. FRET between specific sites on the protein target RNA suggest that both binding regions on FinO interact with the base of SLII.

We show that FinO-promoted duplex formation requires FinO unwinding activity to destabilize RNA stem-loops. Using site-directed mutagenesis, we define a specific region (residues 32-42) associated with this unwinding function. We also show that energy required for unwinding activity is obtained from the binding of FinO to stem-loops.

Acknowledgements

First off, I would like to thank Mark Glover, my supervisor. Mark is an excellent scientist with a passion for discovery. I am grateful that he was willing to take me on and pass on his knowledge and skills. In addition to helping me become a good researcher, Mark has also shown me the fun of science.

I would like to thank our collaborators, Laura Frost and Mike Gubbins, for contributing greatly to the work presented in this thesis. It was a real pleasure to know the two of you and I really enjoyed my trips over to your lab as well as our endless discussions on FinO. Mike, I hope we can hang out at another conference sometime; the one in Texas was a blast! I would also like to thank Tom Kerppola and Vladimir Ramirez-Carrozzi for allowing me to use their facilities in Michigan.

I thank my supervisory committee, Michael James and Lou Agellon, for their guidance through my years of graduate school. I would also like to thank my external examiner, Jack Greenblatt, for reading my thesis and flying from Toronto for my defense.

It has been very enjoyable to work with everyone in the Glover lab including, Dave, David, Jason, Jody, Matt, Nina, Ross, Ruth, Scott and Trevor. I'll miss you all! I would also like to thank everyone in the James lab for their help and for making me feel as if their lab was my lab.

I owe many, many thanks to mom, dad and my brother Chris.

Finally, I will always remember the great friends I had while living in Edmonton for the last few years. Thank you Kathy, Derrick, Sunil, Carey, Martin, Ken and Isabelle, John and Chris.

Table of Contents

Chapter 1. Introduction Page

1.1 Bacterial conjugation 2
1.2 The bacterial conjugation cycle 3
1.3 The F plasmid 5
1.4 Regulation of conjugation 7
1.5 FinO 11
1.6 RNA-protein interactions 12
1.7 RNA binding motifs 14
1.8 HIV RNA binding proteins 15
1.9 HIV-1 nucleocapsid protein NCp7 17
1.10 Regulatory systems involving antisense RNA 18
1.11 Structure and function of Rom 22
1.12 Helicases 24

Chapter 2. Experimental procedures

2.1 Construction of pGEX fusion plasmids 27
2.2 Purification of FinO and FinO-derived proteins 27
2.3 *In vitro* transcription of RNA 31
2.4 Proteolysis 33
2.5 Circular dichroism spectroscopy 34
2.6 Electrophoretic mobility shift assay (EMSA) 34
2.7 Crystallization of FinO(26-186) 36
2.8 Crystallographic data collection and processing 36
2.9 Model building and refinement of FinO(26-186) structure 36
2.10 RNA duplexing assays 37
2.11 Unwinding assay 38
2.12 Protein-RNA cross-linking 39
2.13 GelFRET assay 40
2.14 Mating assays 41

2.15 Northern blot analysis to determine FinP half-life 41

Chapter 3. The FinO repressor of bacterial conjugation contains two RNA binding regions

3.1 Introduction 43
3.2 Results 45
  3.2.1 RNA binding protects N- and C-terminal regions of FinO from proteolysis 45
  3.2.2 Analysis of the structure of FinO by circular dichroism spectroscopy 48
  3.2.3 The N- and C-terminal regions of FinO each specifically bind FinP 51
  3.2.4 FinO binds RNA as a monomer 55
3.3 Discussion 57

Chapter 4. Crystal structure of the bacterial conjugation repressor FinO

4.1 Introduction 61
4.2 Results 62
  4.2.1 Structure determination 62
  4.2.2 Structure description 64
  4.2.3 FinO conformational flexibility 64
  4.2.4 FinO-RNA interactions 71
  4.2.5 FinP-*traJ* RNA association 72
4.3 Discussion 73

Chapter 5. Probing FinO – FinP RNA interactions by site-directed protein-RNA crosslinking and gelFRET

5.1 Introduction 76
5.2 Results and discussion 80
  5.2.1 Site-specific FinO-SLII RNA cross-linking 80
  5.2.2 Probing FinO-RNA architecture using gelFRET 83

Chapter 6. The FinO repressor of bacterial conjugation is an RNA chaperone that facilitates sense-antisense RNA interactions

6.1 Introduction 89
6.2 Results 93
6.2.1 FinO can unwind dsRNA 93
6.2.2 RNA unwinding activity is associated with the N-terminal region of FinO 97
6.2.3 FinO utilizes its RNA-binding energy to unwind RNA 99
6.2.4 Residues 26 to 44 of FinO function to promote FinP-*traJ* RNA duplex formation 103
6.2.5 The RNA unwinding and duplexing activities of FinO are essential for the repression of bacterial conjugation 105
6.3 Discussion 107

Chapter 7 Summary and Future Perspective 112

Chapter 8 References 119

List of Tables

Chapter 2 Page

2.1 List of all primers used in the construction of FinO derrivatives 28
2.2 List of DNA templates used in synthesis of RNA substrates 32

Chapter 3

3.1 Apparent Ka of the FinO deletion proteins binding to FinP 54

Chapter 4

4.1 X-ray data collection, phasing and refinement statistics 66

Chapter 6

6.1 Unwinding rates, duplexing rates and mating efficiencies for FinO and FinO derivatives 102

List of Figures

Chapter 1

Page

1.1 The bacterial conjugative cycle 4
1.2 F plasmid conjugative transfer genes 6
1.3 FinOP regulatory system 8
1.4 Molecular details of FinO mediated conjugative inhibition 9
1.5 Primary sequences and secondary structures of FinP and the 5' UTR of *traJ* mRNA 10
1.6 Structures of NCp7 bound to target stem-loops 16
1.7 ColE1 replication regulation by RNAI, RNAII and Rom 20
1.8 CopT/CopA regulatory system of IncFII-like plasmids 21
1.9 Model of Rom bound to a kissing complex 23

Chapter 3

3.1 Primary sequence of FinO and FinP 44
3.2 N- and C-terminal regions of FinO are protected against trypsin digestion by FinP RNA 47
3.3 Denaturing PAGE of purified FinO and FinO fragments 49
3.4 Circular dichroism spectroscopy of FinO and FinO fragments 50
3.5 Measurement of the ability of FinO deletion proteins to bind FinP RNA by EMSA 52
3.6 EMSA demonstrates that FinO binds FinP as a monomer 56

Chapter 4

4.1 Stereo view of electron density maps around residues 152-156 63
4.2 The primary sequence and secondary structure elements of FinO 65
4.3 Ribbon diagram of FinO 67
4.4 Molecular surface representation of FinO, colored by electrostatic potential 68

4.5 Trace of main chain Cα atoms and their B-factors 69
4.6 FinO-RNA interactions 70
4.7 Duplexing assays for FinO(1-186) and FinO(26-186) 74

Chapter 5

5.1 Schematic representation of SLII-based RNAs used in this study 77
5.2 Ribbons diagrams of FinO 78
5.3 Schematic representations of FinO 79
5.4 Site specific cross-linking of FinO and SLII 82
5.5 Probing the interactions of FinO bound to a target RNA by gelFRET 85
5.6 Labeling of FinO with Texas Red leads to changes in the mobilities of several of the FinO-SLII complexes 86

Chapter 6

6.1 The sequence and secondary structure of RNA molecules used in this study 90
6.2 FinO secondary structure and the N-terminal sequence 91
6.3 Schematic diagram of the unwinding assay 94
6.4 N-terminus of FinO is required for unwinding activity 95
6.5 Unwinding activity is highly dependent on protein concentration 96
6.6 Unwinding efficiencies of N-terminal double alanine point mutants 98
6.7 Unwinding efficiency of N-terminal single alanine point mutants 100
6.8 Representative EMSAs for FinO, FinO(26-186) or FinO(45-186) binding to SLII 101
6.9 Duplexing assays with FinO, FinO(26-186) or FinO(45-186) 104
6.10 Duplexing efficiencies of N-terminal mutants 106
6.11 N-terminus of FinO is not required for FinP stabilization 108

Chapter 7

7.1 Model for how FinO might promote duplex formation between complementary stem-loop structures by unwinding the base of the stem 115
7.2 Model for how FinO might promote duplex formation between complementary stem-loop structures by unwinding the stem region below the loop 116

List of Abbreviations

ATP	adenosine 5'-triphosphate
Å	Ångstrom
APA-Br	azidophenacyl bromide
βME	β-mercaptoethanol
BSA	bovine serum albumin
CD	circular dichroism
CTP	cytidine 5'-triphosphate
DNA	deoxyribonucleic acid
DIS	dimerization initiation site
DMSO	dimethyl sulfoxide
DTNB	dithionitrobenzoate
DDT	dithiothreitol
dsDNA	double-stranded DNA
dsRNA	double-stranded RNA
dsRBD	double-stranded RNA binding domain
E. coli	*Escherichia coli*
EMSA	electrophoretic mobility shift assays
EDTA	(ethylenedinitrilo)-tetraacetic acid
FPLC	fast performance liquid chromatography
Fin	fertility inhibition
FRET	fluorescence resonance energy transfer
GST	glutathione-S-transferase
GTP	guanosine 5'-triphosphate
HPLC	high performance liquid chromatography
HIV	human immune deficiency virus
Inc	incompatability
IPTG	isopropyl β-D-thiogalactopyranoside
LB	Luria Bertani
K_a	equilibrium association constant
k_1	first order rate constant
k_2	second order rate constant
kDa	kiloDaltons

MS	mass spectrometry
mRNA	messenger RNA
MES	(2-[N-morpholino]ethanesulfonic acid)
MAD	multiple wave-length anomalous dispersion
NC	nucleocapsid
NMR	nuclear magnetic resonance
ORF	open reading frame
PEG	polyethylene glycol
PMSF	phenylmethyl sulfonly fluoride
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
P	promoter
RRM	arginine rich motif
RNA	ribonucleic acid
RNP	ribonucleoprotein
RBS	ribosomal binding site
SDS	sodium dodecylsulfate
SL	stem-loop
TR	Texas Red
tra	transfer
Tris	tris(hydroxymethyl)aminomethane
TBE	Tris-borate-EDTA
UV	ultraviolet
UTR	untranslated region
UTP	uridine 5’-triphosphate

Chapter 1

Introduction

1.1 Bacterial conjugation

Bacterial conjugation is the process by which genetic material is transferred between two bacterial cells in close contact with each other. Lederberg and Tatum first discovered this process in 1946 when they observed that two different auxotrophic populations of *Escherichia coli* K-12 could be mixed together and yield recombinants having the same phenotype as wild-type *E. coli* K-12. Passage of a soluble factor between cells was ruled out and it was suggested that cell fusion precedes the horizontal transfer of DNA. Subsequently, Davis (1950) established that cell-cell contact was indeed necessary for mating to occur between bacteria.

Independent studies by Lederberg and Cavalli (1952) and Hayes (1953) identified an infectious agent responsible for DNA transfer during conjugation and named it Fertility, or F factor. It was later shown that this factor was a double-stranded, circular DNA and was termed the F plasmid.

The last 40 years have lead to the discovery of many conjugative plasmids and it is now established that conjugation is a general phenomenon used by bacteria to rapidly disseminate genetic material across species and genera (Silver & Bostian, 1993). Boyd and Hartl (1997, 1998) showed that horizontal transfer of the F plasmids between *Escherichia coli* and *Salmonella enterica* has occurred in the recent evolutionary past and they suggest that F plasmids encode for the conjugative machinery required for transfer of virulence factors between the different *Salmonella* sub-species. Conjugation is the primary route for bacterial cells to pass determinants of antibiotic resistance to one another and this has important medical implications, since the number of drugs that can be used to fight bacterial infections becomes increasingly limited as bacterial strains add additional drug resistance genes to their genetic repertoire. In addition many of the plasmids characterized, contain resistance determinants to multiple antibiotics (Couturier, 1988). A microbe picking up one of these plasmids can rapidly increase its spectrum of resistance. One important family of plasmids whose members often encode multiple drug resistant genes is the R plasmids. These plasmids are largely responsible for the spread of resistance in strains of *E. coli* and were first isolated in Japan, where they were shown to be the causative agent of antibiotic-resistant dysentery in Japan following World War II (Watanabe & Fukasawa, 1961).

Plasmids are classified into various groups and subgroups based on their conjugative, replicative and regulatory components. Plasmids with the same conjugative systems are placed into the same group. These groups are then divided into subgroups that share the same replication systems and cannot stably coexist within a growing bacterial population (Datta, 1975; Austin & Nordström, 1990). The subdivisions are therefore referred to as incompatibility groups. The F plasmid, along with ColV and R453, belong to the group IncF, sub-division I, or simply group IncFI (Ippen-Ihler & Skurray, 1993). The taxonomy is a bit misleading, since the main groups also contain the abbreviation for the term incompatability (e.g. IncF), but only plasmids within a particular subgroup (e.g. IncFI) are incompatible.

1.2 The bacterial conjugation cycle

The majority of proteins functioning in conjugative transfer are encoded by the donor plasmid, therefore the mechanism of conjugation is dependent on the type of plasmid being transferred. The conjugative model described below is specific for the mechanism encoded by a large group of plasmids, referred to as the F-like plasmids (Figure 1.1), with the prototypic model originating from the study of the F-plasmid. For a review of F-plasmid conjugation see Firth *et al.* (1996).

The conjugation process requires the expression of a long, filamentous (1 to 4 µm) cell surface pilus that is used to recognize and associate with suitable acceptor cells (Frost et al., 1994). After stable attachment of a pilus tip to specific receptors on the recipient cell, an unknown signal is sent back to the host cell to initiate the retraction of the pilus through depolymerization of the pilin subunits (Novotny & Fives-Taylor, 1974). With pilus retraction, the donor and recipient cells are brought into close contact and outer-membrane proteins between the mating pair can associate to stabilize the pair (Achtman *et al.*, 1977) and allow for a higher frequency of mating (Miki, 1978; Firth & Skurray, 1992). Through an unknown signal, the plasmid-encoded relaxase, TraI, is activated and nicks the positive strand of the plasmid at the origin of transfer (oriT) locus (Willetts, 1972; Thompson *et al.*, 1984). Subsequently, the nicked strand is transferred to the recipient, in the 5' to 3' direction (Ohki & Tomizawa, 1968; Rupp & Ihler, 1968), concomitant with the synthesis of the complementary strands in

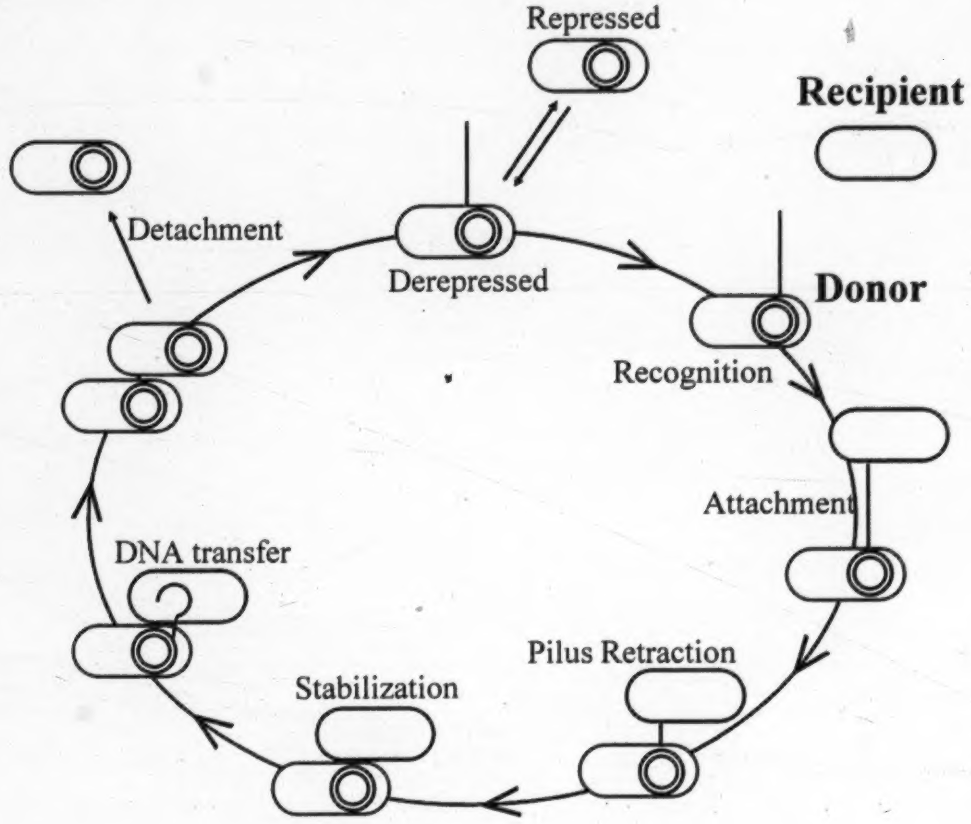


Figure 1.1: The bacterial conjugative cycle. Donors are shown in red, while the recipients are colored black. Release of conjugative inhibition allows the donor to enter the conjugative cycle and transfer the F-plasmid to a recipient cell. Adapted from Jerome (1999).

both the host and recipient (Sarathy & Siddiqi, 1973; Ohki & Tomizawa, 1968; Vapnek & Rupp, 1970). Replication of the host plasmid terminates at the *oriT* site and the positive strand is re-ligated (Wilkins & Lanka, 1993). In the recipient cell, the two strands of the new resident plasmid also have to be ligated into a circular double stranded DNA. The recipient now produces plasmid-encoded exclusion factors to reduce the probability of homosexual mating. These factors are highly specific and will only inhibit the entry of related plasmids in addition to other F-plasmids (Willetts & Maule, 1974). The F plasmid encodes the exclusion factors TraT and TraS. TraT functions as a cell surface exclusion factor, while TraS is an entry exclusion factor, which prevents the transfer of DNA between donors harboring like-plasmids (Achtman *et al.*, 1977). In the final stage of the conjugation, the mating pair separates and the newly infected cell, as well as the donor, can initiate another round of the conjugative cycle.

1.3 The F Plasmid

The proteins required for conjugation of F plasmids are encoded within the 33 kb transfer (*tra*) region encompassing one-third of the entire F plasmid sequence (Figure 1.2). The majority of the *tra* region's 36 open reading frames (ORFs) are situated within the multi-cistronic *tra* operon, with only the three regulatory gene products TraM and TraJ and FinP being encoded outside of this operon.

The *tra* region gene products are classified based on their participation in one of the following four stages of bacterial conjugation: 1) regulation, 2) pilus synthesis and assembly, 3) aggregate stability, or 4) signaling, origin nicking unwinding, and transport. For the complete *tra* region sequence, and a list of known gene product functions, see Frost *et al.* (1994). The work presented within this thesis deals with regulatory aspects of conjugation and the next sections will give an overview of the current knowledge of regulation. A detailed review on the other stages of bacterial conjugation is given in *Bacterial Conjugation* (1993).

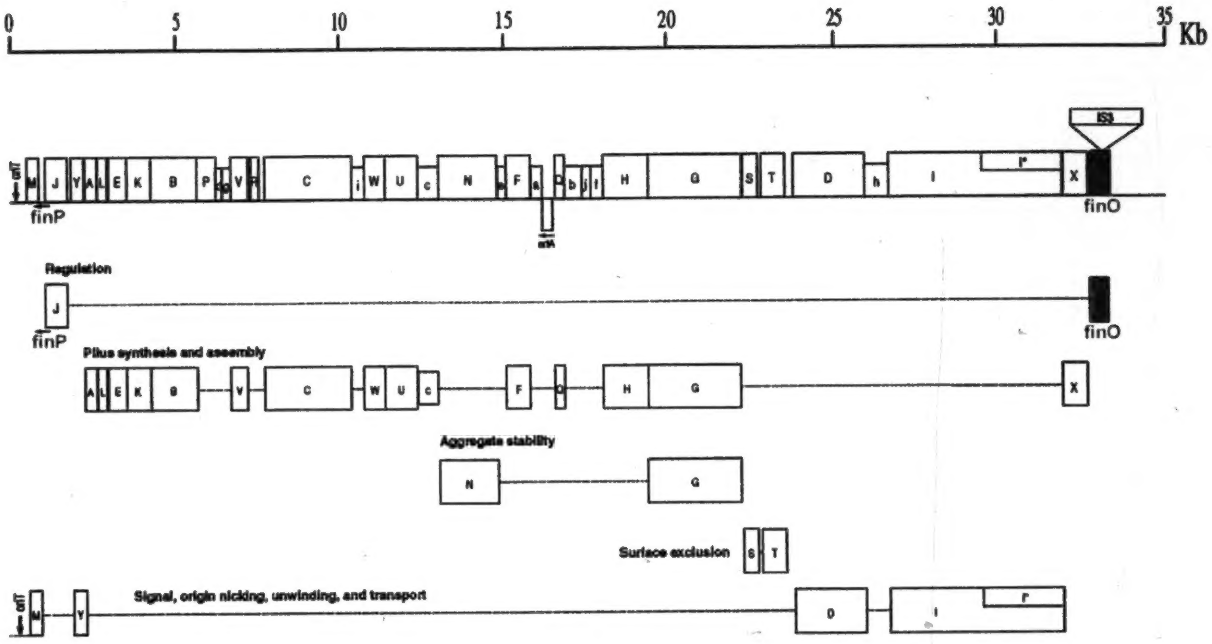


Figure 1.2: F plasmid conjugative transfer genes. The transfer genes are divided into their respective groups at the bottom of the figure, while the length of the transfer region is shown at the top. FinO and finP are highlighted in red. Adapted from Frost *et al*. (1994).

1.4 Regulation of conjugation

The transcriptional regulation of conjugation is orchestrated around the 5' end of the *tra* region (Figure 1.3). For conjugation to occur, the activation of both the *tra* operon promoter (P_Y) and the *traM* promoter (P_M) is required. This allows for the expression of all essential transfer proteins. Transcription from P_Y requires the activity of TraJ, a plasmid-encoded transcriptional activator that contains a DNA-binding helix-loop-helix motif. TraJ has not yet been shown to directly contact P_Y. It is thought that maximum activation of the P_Y promoter requires TraJ to function in a nucleoprotein complex containing the host-encoded activator protein ArcA, as well as the first gene product of the *tra* operon, TraY (Gaudin & Silverman, 1993). In addition to autoregulating transcription of its own mRNA, TraY also binds the P_M promoter and upregulates production of TraM, a protein that binds to the origin of transfer (*oriT*) and is required for plasmid replication (Penfold *et al.*, 1996). As the levels of TraM increase, a negative feedback loop is initiated, where TraM binds to its own promoter region and inhibits the synthesis of *traM* mRNA (Schwab *et al.*, 1993).

The RNA of the transcriptional activator TraJ contains a 105 base 5' untranslated leader region (UTR) that folds into three stem-loop (SL) structures numbered from the 3' end, as Ic, IIc and III (Figure 1.5) (van Biesen *et al.*, 1993). Translation of TraJ requires binding of the ribosome to an eight-base ribosomal binding site (RBS) situated in the loop and stem of SLIc.

F and F-like plasmids have evolved an elegant two-component system for fertility inhibition that involves repression of TraJ translation by occluding the *traJ* RBS. This system has been termed FinOP and consists of the antisense RNA component FinP and the RNA binding protein FinO (Figure 1.4). The ~79 nucleotides that constitute FinP are transcribed in the opposite direction to *traJ*, from a promoter region situated within the leader sequence of *traJ* mRNA. FinP folds into two stem-loop structures (SLI and SLII) that are complementary to SLIc and SLIIc of the *traJ* leader sequence (Mullineaux & Willetts, 1985). van Biesen *et al.*, (1993) demonstrated that FinP and the *traJ* mRNA can form a duplex *in vitro* and it was suggested the formation of this duplex masks the RBS and prevents TraJ translation.

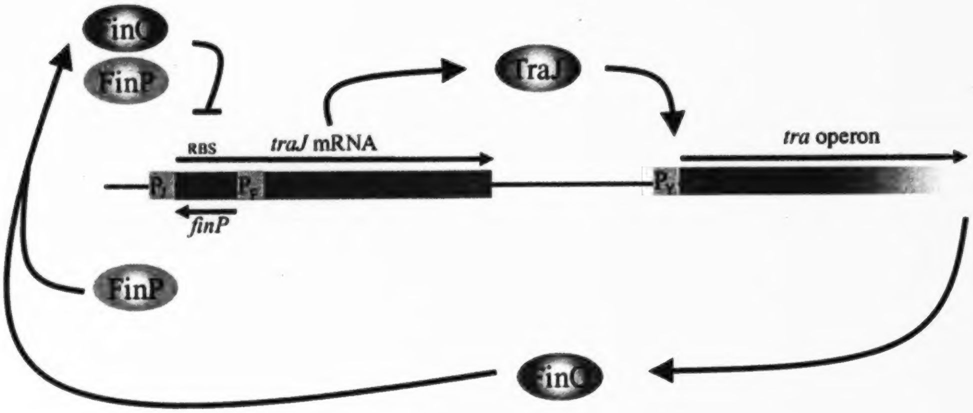


Figure 1.3: FinOP regulatory system. Bacterial conjugation is inhibited by the two-component FinOP system. The protein FinO promotes the association of the antisense RNA FinP with the 5' UTR of *traJ* mRNA. This association occludes the *traJ* mRNA RBS and inhibits TraJ translation. TraJ is the transcriptional activator for the *tra* operon and without TraJ the tra operon gene products are not expressed and conjugation is repressed.

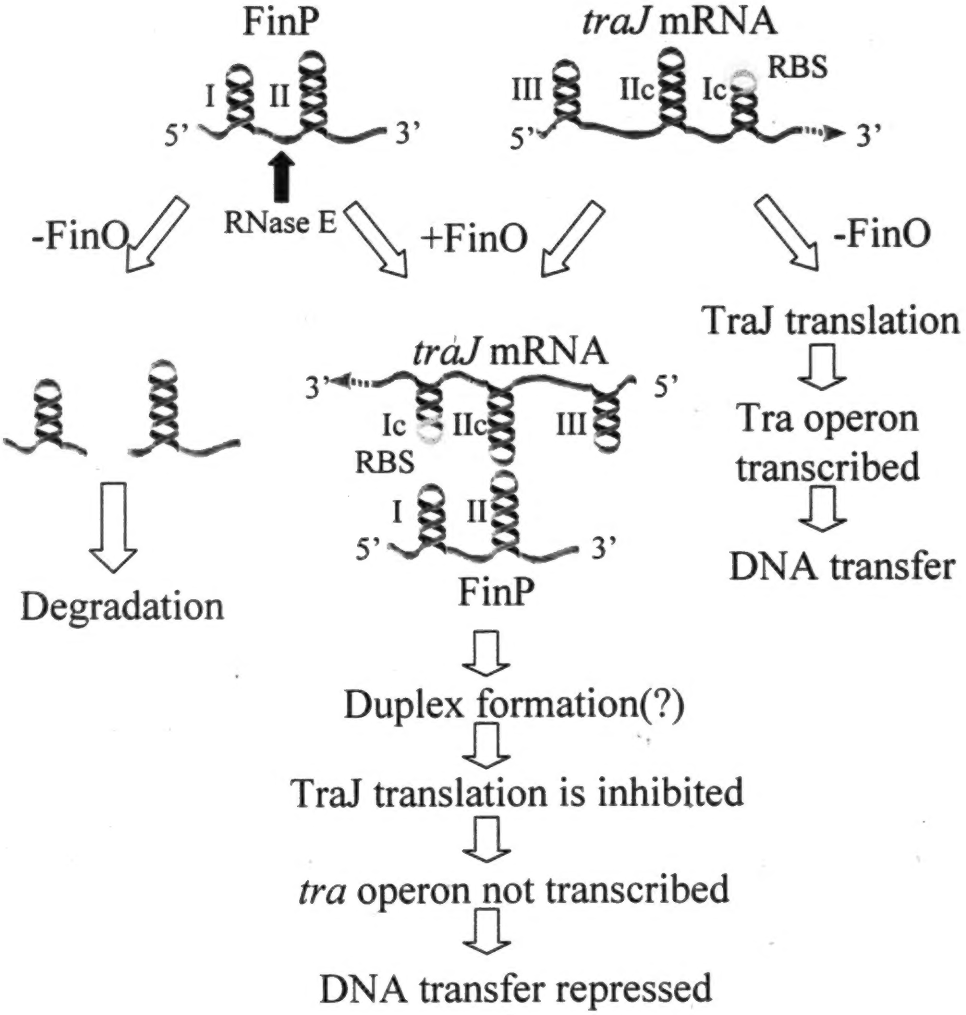
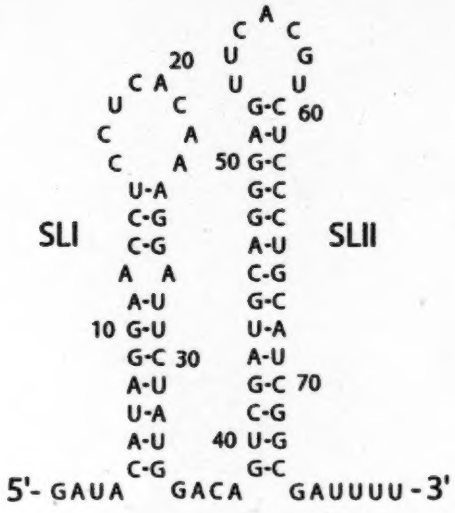
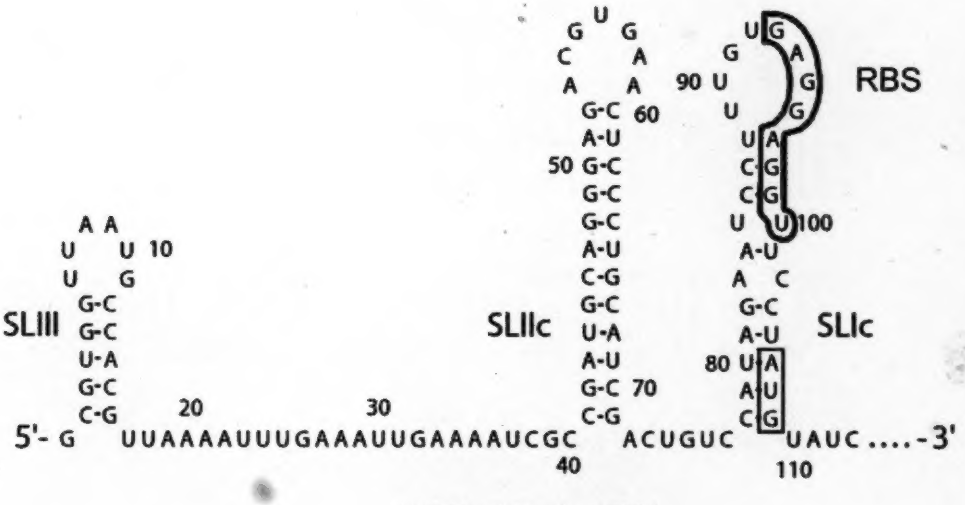


Figure 1.4: Molecular details of FinO mediated conjugative inhibition. FinO binds to both FinP and *traJ* mRNA, protecting FinP from Rnase E degradation and possibly promoting duplex formation between FinP and *traJ* mRNA. Association of FinP with *traJ* mRNA blocks a RBS (shown as a yellow circle) on *traJ* mRNA and results in conjugative inhibition. Adapted from Lori Jerome, 1999.



FinP



5' UTR of *traJ* mRNA

Figure 1.5: Primary sequences and secondary structures of FinP and the 5' UTR of *traJ* mRNA. The sequences are numbered from their 5' ends. The stem-loop names are given, the start codon of *traJ* is boxed as is the ribosomal binding site (RBS).

For FinP, as is the case in several other sense-antisense regulatory systems, the first critical interactions in duplex formation involve the formation of a ‘kissing’ complex between complementary loops (i.e. two or three intermolecular base pairs forming between the loops), as is described in more detail in section 1.10 (Koraimann *et al.*, 1996). Since the loop sequences are highly variable between different *finP* alleles of the IncF plasmids, FinP from one plasmid is unable to associate with *traJ* mRNA from a different co-resident plasmid (Finlay *et al.*, 1986; Willetts & Maule, 1986; Koraimann *et al.*, 1991; Koraimann *et al.*, 1996). Therefore, FinP cannot block translation of TraJ from a different plasmid and this defines FinP plasmid specificity.

FinP works in concert with FinO to inhibit conjugation. FinO serves two purposes: to enhance duplex formation between FinP and *traJ* mRNA (van Biesen & Frost, 1994), and to stabilize FinP from cellular RNaseE (Jerome *et al.*, 1999). Unprotected FinP is rapidly depleted from the cell by RNaseE cleavage at a site between SLI and SLII. The presence of FinO was shown to increase the *in vivo* half-life of FinP by over 10-fold.

1.5 FinO

FinO is a highly basic protein (estimated pI = 9.7) that has a mass of approximately 21.2 kDa and consists of 186 amino acids. The protein is encoded at the 3’ end of the *tra* operon (Cheah & Skurray, 1986; Yoshioka *et al.*, 1987; Yoshioka *et al.*, 1990). The model IncF plasmid, F, has an insertional (IS3) element disrupting the *finO* gene and requires a functional *finO* gene supplemented *in trans* for conjugative repression. Unlike FinP, FinO is plasmid non-specific; one *finO* allele is able to substitute for mutated *finO* alleles in various IncF plasmids (Finnegan & Willetts, 1971). The *finO* alleles are categorized into two groups, referred to as, Type I and II (Willetts & Maule, 1986). Although these two *finO* genes are highly conserved, the level of FinO expression is much greater in Type I as a result of mRNA stabilization by the presence of an additional open reading frame (*orfC*) on the mRNA that is 5’ to *finO* (van Biesen & Frost, 1992). This increased mRNA stability accounts for the 5- to 50- fold increase in conjugative repression observed by Type I over Type II (Willetts & Maule, 1986).

Studies by van Biesen & Frost (1994) revealed that FinO binds both FinP and *traJ* mRNA and that SLII is the minimal target to which FinO can bind with high affinity. Later work performed by Jerome *et al.* (1999) on numerous stem-loop constructs further defined the minimal binding target of FinO as stem-loop structures that have a 3' single-stranded tail. The length of the stem and the 3' tail were found to be important determinants for high affinity binding. These results helped explain why FinO does not bind as tightly to SLI, which has a shorter 3' tail than SLII (Figure 1.5). Furthermore, binding was not affected by the sequence of the RNA or by the size of the loop and was not dramatically affected by bulges in the stem.

Although FinO protects FinP from RNase E degradation and promotes duplex formation between FinP and *traJ* mRNA *in vitro*, the *in vivo* significance of these individual functions and the molecular mechanism by which FinO accomplishes these feats remain to be determined. In this thesis I will present results that begin to shed light on the molecular mechanism by which FinO helps to promote duplex formation and, in addition, studies are presented that show for the first time that FinO-mediated RNA duplexing is required for inhibition of conjugation.

1.6 RNA-protein interactions

Cellular RNA molecules generally contain both single stranded regions and stem-loop structures of various lengths and sequences. Stem-loops often form the binding sites for proteins, as is the case with FinO. The major grooves of the RNA stems adopt an A-form helix that is deeper, narrower and less accessible to the protein side chains than the major groove of double stranded DNA. To achieve specificity, many RNA binding proteins interact with the more accessible bases of single stranded RNA, such as out turned bases in loops and the tails at the base of the stems. Additionally, bulges, kinks and mismatches can distort the grooves of double helical RNA to increase protein accessibility to the bases (Draper, 1999).

In contrast to the major groove of either RNA or DNA, the single stranded RNA regions are much more accessible to protein main-chain amide and carbonyl groups (Allers & Shamoo, 2001). As a result, only 10% of the hydrogen bonds occurring between proteins and DNA bases involve the peptide backbone, whereas RNA binding proteins use the backbone amides and carbonyls 33% of the time. Analysis of 45 different RNA/protein structures revealed that the protein backbone is the most frequently used determinant for hydrogen bonding to RNA bases (33%) and the 2' hydroxyl group of ribose (39%). The guanidinium group of arginine also makes large numbers of hydrogen bonds with the RNA bases (26%) and the ribose –OH group (13%). When it comes to interacting with the backbone phosphodiester group, arginines and lysines are the best-suited groups, since they have the ability to form both electrostatic interactions and hydrogen bonds.

RNA can also adopt higher ordered structures and in this case, specificity can be achieved by recognition of the specific tertiary elements. This type of interaction was observed in the crystal structure of the ribosomal L11-rRNA complex. The L11 protein makes contacts with, and also stabilizes, a critical RNA tertiary interaction involving an adenosine that packs against a stem structure (Conn *et al.*, 1999).

RNA-protein complexes often contain stacking interactions between aromatic residues in the protein and RNA bases. In the U1A-RNA hairpin complex, the phenyl group of a tyrosine residue stacks with the pyrimidine ring of a cytosine in the loop and stabilizes the complex (Oubridge *et al.*, 1994). Removal of this tyrosine completely abolished binding (Jessen *et al.*, 1991). In this same structure, two other bases in the loop are sandwiched between a phenylalanine and an aspartate to form a four-element stack. The crystal structure of the aspartyl-tRNA synthetase complexed with its cognate tRNA reveals the presence of a phenyl ring stacked against a uracil in the anti-codon loop (Cavarelli *et al.*, 1993). Based on the crystal structure of NS3 helicase from the hepatitis C virus bound to a DNA substrate (Kim *et al.*, 1998), it was suggested that the intercalation of a highly conserved tryptophan residue between bases is of functional significance in unwinding of cognate double stranded RNA. The binding of the phage λ N protein to its target stem-loop RNA requires a tryptophan residue; substitution to any other residue decreased the binding affinity (Su *et al.*, 1997). NMR structures of this complex reveal the stacking interaction of the tryptophan with a base in the RNA tetraloop (Legault *et al.*, 1998; Scharpf *et al.*, 2000).

1.7 RNA binding motifs

Proteins use a variety of RNA binding motifs to recognize RNA (Burd & Dreyfuss, 1994). The best characterized, and most common RNA-binding motif is referred to as either the RNA recognition motif (RRM) or the ribonucleoprotein (RNP) motif. The RNP motif contains two short sequences, RNP1 and RNP2, that are highly conserved and make a general RNA-binding module on RNP-domains. The structures of several RNP domains have been determined, including the spliceosomal U1A protein and U2B"-U2A' protein complex bound to cognate RNA stem-loop structures (Oubridge *et al.*, 1994; Price *et al.*, 1998). Although all RNP-domains have nearly identical folds, there is significant variation in the amino acid sequence. This variation allows each RNP to form a unique network of interactions with the target RNA and helps confer binding specificity (Price *et al.*, 1998).

The arginine rich motif (ARM) is another common motif used by proteins to bind RNA. The ARM consists of a few residues that contain a propensity for arginines. Several RNA binding proteins associate with their target RNA through an ARM that can be separated from the rest of the protein and still bind tightly to the RNA. The relatively small size of complexes formed between stem-loops and their minimal binding peptides has facilitated the structural determination by NMR of many of these complexes (Pulisi *et al.*, 1995; Battiste *et al.*, 1996; Legault *et al.*, 1998; Cai *et al.*, 1998; Faber *et al.*, 2001).

The RGG box is another motif that contains numerous arginines. This motif is characterized by the presence of multiple (up to 18), closely spaced, arginine-glycine-glycine (RGG) repeats (Burd & Dreyfuss, 1994).

The double-stranded RNA-binding domain (dsRBD) is a conserved motif found as one or more copies in several proteins that bind to regions of dsRNA in a sequence non-specific manner. These motifs contain many basic residues at conserved positions and are able to discriminate between dsRNA and dsDNA. Several studies, including the structural determination of a dsRBD in complex with dsRNA, have revealed that the specificity for dsRNA over dsDNA results from the formation of a network of interactions with dsRNA that is dependent on the unique geometry of A-form RNA (Ryter & Schultz, 1998; Bevilacqua & Cech, 1996). Consequently, many of these interactions could not occur with canonical B-form DNA.

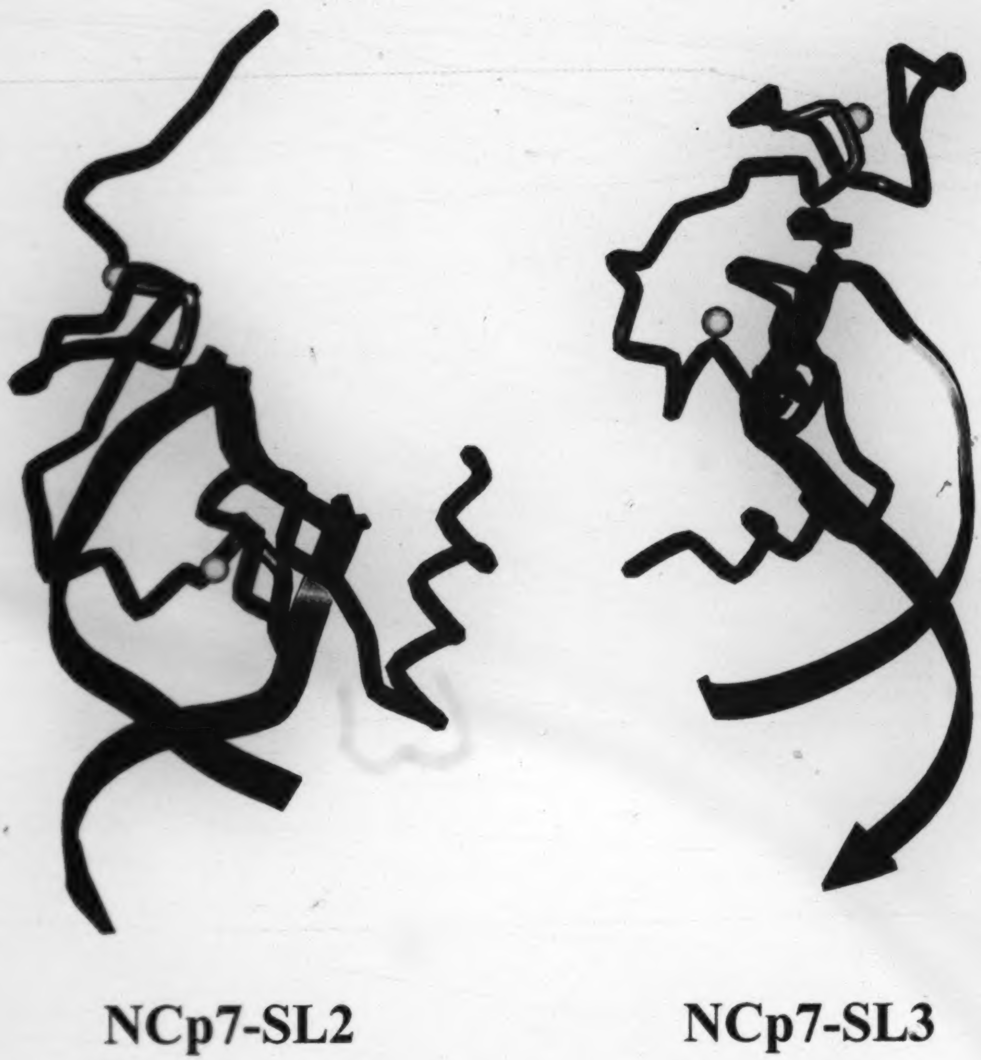
The zinc knuckle is another motif found in RNA binding proteins. This motif contains a characteristic fold that is structured around a zinc atom coordinated by cysteines and histidines. In the NMR structure of the HIV nucleocapsid protein NCp7, the zinc knuckles are shown to make critical contacts with solvent exposed guanines in the loops of hairpins (Figure 1.6) (Amarasinghe *et al.*, 2000; De Guzman *et al.*, 1998). NCp7 is described in more detail later on.

1.8 HIV RNA binding proteins

HIV RNA transcripts are structured into subdomains that are active participants in many of the biochemical processes observed in the viral life cycle (Frankel & Young, 1998). These subdomains contain hairpins that interact with target proteins and contribute to viral processes including regulation of transcription, nuclear export, viral packaging and maturation. Binding of the transcriptional activator of HIV (Tat) to its cognate stem-loop target (TAR) promotes transcriptional antitermination and allows efficient synthesis of full length genomic RNA (Karn, 1999). Rev is another HIV encoded protein and is required for control of viral gene expression. Binding of Rev to the RRE stem-loop structure is involved in regulation of viral RNA splicing and the subsequent export of RNA from the nucleus to the cytoplasm (Pollard & Malim, 1998).

The HIV genome consists of two RNA molecules that are associated (or dimerized) prior to packaging into the viral particle. Dimerization initially occurs through the formation of a tight kissing complex between a stem-loop at the dimerization initiation site (DIS) of one RNA molecule and the complementary stem-loop of the other RNA (Paillart *et al.*, 1996; Laughrea *et al.*, 1997). Following release of the virus from the host cell, maturation of the viral particle is required and this involves the conversion of the kissing complex to an extended duplex between the two complementary stem-loops (Parslow *et al.*, 1996). The nucleocapsid protein, NCp7, assists in this dimerization process by recognizing the kissing complex intermediate and promoting duplex formation (Feng *et al.*, 1996; Feng *et al.*, 1995; Muriaux *et al.*, 1996).

The DIS comprises one of the four hairpins in the RNA ψ-packaging signal required for viral assembly (Lever *et al.*, 1989). NCp7, as part of the larger gag



NCp7-SL2

NCp7-SL3

Figure 1.6: Structures of NCp7 bound to target stem-loops. The solution structures of NCp7 in complex with SL2 and SL3 were determined by NMR spectroscopy (Amarahinghe *et al.*, 2000; Guzman *et al.*, 1998). Zinc knuckles (purple) specifically recognize guanines (green) that are expelled from the loops. Zinc ions are indicated in yellow. Adapted from Amarahinghe *et al.*, 2000 and Guzman *et al.*, 1998.

polyprotein, also interacts with these hairpins during Gag-mediated transport of genomic RNA to the cell surface for packaging (reviewed in Turner & Summers, 1999).

1.9 HIV-1 nucleocapsid protein NCp7

NCp7 is a small 55 amino acid protein that contains two zinc finger motifs flanked on either side by highly basic regions. This protein has many functional roles in the retroviral life cycle that require binding of this protein either specifically to RNA stem-loop structures or non-specifically to single stranded nucleic acid regions. NCp7 activities include: 1) associating with the packaging signal of HIV and transporting the genomic RNA to the cell surface for encapsidation, 2) initiating synthesis of a DNA copy of the viral genome by annealing cellular tRNA^{lys} to the transcriptional start site, 3) dimerization of the viral RNA during the maturation process, and 4) reducing reverse transcriptase pausing during HIV DNA synthesis by destabilizing secondary structures within the RNA template (reviewed in Rein *et al.*, 1998). As described above, the role of NCp7 (as part of the gag polyprotein) in viral encapsidation is to recognize specific stem-loop structures in the ψ-packaging signal. All the other processes involve NCp7 nucleic acid chaperone activity, where preformed base-pairs are destabilized to either, 1) specifically remove secondary structures in preventing transcriptional pausing (Wu *et al.*, 1996; Ji *et al.*, 1996) or 2), to remove intramolecular stem-loops and promote formation of lower energy duplexes with the maximum number of base-pairs, as is observed in the maturation process (Feng *et al.*, 1996) and in the unwinding and subsequent association of tRNA with the transcriptional start site (Khan *et al.*, 1992; Lapadat-Tapolsky *et al.*, 1995). The solution structures of NCp7 bound to either SL2 or SL3 of the psi-packaging signal reveal that the zinc fingers make specific contacts with the loops, while the N-terminal basic region interacts with the stem (describe in Figure 1.6) (Amarasinghe *et al.*, 2000; De Guzman *et al.*, 1998). Although binding of hairpins by NCp7 requires the zinc finger motifs, the associated chaperone activity does not require the zinc fingers and is only dependent on the basic regions surrounding the zinc fingers (Takahashi *et al.*, 2001).

1.10 Regulatory systems involving antisense RNA

Nature has implemented antisense RNA in controlling a wide range of biological processes, ranging from DNA replication of the ColE1 plasmid in bacteria to regulation of gene expression in both prokaryotes and eukaryotes (reviewed in Wagner & Simons, 1994; Knee & Murphy, 1997). Several natural antisense regulatory systems have been well characterized in bacteria and although studies on eukaryotic antisense systems have lagged behind those in bacteria, mounting evidence suggests that many systems in eukaryotes are, at least in part, controlled by antisense RNA. Genomic characterization has revealed that antisense regulation is likely a general phenomenon adopted by all species, from archaebacteria to humans (Stolt & Zillig, 1993; Merino *et al.*, 1994).

Generally, antisense regulatory systems in bacteria involve the melting of base-paired regions in stem structures of the antisense RNA and the subsequent hybridization of these regions with complementary stem-loop regions in the target (Wagner & Simons, 1994). Dimerization between antisense RNA and its target requires an initial nucleation event to occur between complementary loops (Eguchi *et al.*, 1991). Loops involved in this primary recognition event typically range in size from four to seven nucleotides (Hjalt & Wagner, 1992). Nucleation involves the reversible base pairing between one or two sets of complementary loops (Houssier & Grosjean, 1985; Yoon *et al.*, 1975). The stability of this initial kissing complex is dependent on the bases involved in the pairing interactions and is enhanced by stacking interactions with additional nucleotides on the 3' side of the loop (Yoon *et al.*, 1975; Eisinger & Sphahr, 1973; Pongs *et al.*, 1973; Freier & Tinoco, 1975). The sequence of many loops involved in kissing complex formation, including SLI and SLII from FinP, contain U-turn motifs to help preorder the bases in a manner suitable for rapid and stable association of the complementary loops (Franch *et al.*, 1999). U-turn motif's have the signature sequence 5'-U/G_1-N_2-R_3-3' (N=G, A, C, U; R=G, A) (Franch *et al.*, 1999; Gutell *et al.*, 2000). As observed in several structures, a distinguishing feature of U-turns is the interaction of the U/G_1 base with the phosphate backbone of the third position purine (Pley *et al.*, 1994; Jucker & Pardi, 1995; Stallings & Moore, 1997; Puglisi & Puglisi, 1998; Westhof *et al.*, 1988). This interaction produces a sharp reversal of the RNA backbone after the U/G_1 base that helps to close the loop and also to increase the solvent exposure of the nucleotides on the 3' side of the U/G_1 base.

The solvent accessibility of these bases helps to facilitate interactions with proteins and complementary RNA loops by reducing the entropic cost encountered when binding flexible loops. In ribosome assembly, initial interactions between U-turn motifs and complementary sequences are thought to act as anchors for subsequent tertiary interactions (Cate *et al.*, 1999). Bulges in stem structures and single-stranded tails on either end of the stems also enhance duplex formation by increasing the number of nucleotides that can form base pairs prior to stem unwinding (Hjalt & Wagner, 1995; Kolb *et al.*, 2001).

Two of the better-characterized antisense regulatory systems involve the inhibition of replication of ColE1 and IncFII plasmids. Although both systems function to shut off replication, their mechanisms are strikingly different. Replication inhibition in the IncFII-like plasmids involves binding of the antisense regulator CopA to the target RNA CopT (Figure 1.7) (Womble *et al.*, 1984). CopT is an mRNA that encodes two proteins: RepA, a cis-acting protein that functions to promote the initiation of plasmid replication (Masai *et al.*, 1983), and Tap, a small translational activator peptide encoded upstream of *repA* (Wagner *et al.*, 1987). The *repA* RBS is situated within a region of CopT that can form a stem loop structure. Formation of this stem-loop occludes the ribosomal binding site and inhibits the translation of RepA (Masai & Arai, 1988). Removal of the RepA translation block is directly coupled to the translation of Tap (Blomberg *et al.*, 1992). The RBS of *tap* is not occluded, so ribosomes can efficiently load and translate Tap. As the ribosome reaches the 3' end of the tap ORF, they disrupt the inhibitory stem-loop structure of CopT, allowing ribosomes to reload at the repA RBS and initiate translation of RepA in a process referred to as translational coupling. The antisense regulator CopA is complementary to the translational start site of *tap* and, upon association with CopT, blocks the RBS of *tap* and, in turn, prevents translation of RepA.

In ColE1 plasmids, hybridization of RNAII to its coding DNA strand, followed by RnaseH cleavage of RNAII at a specific site, yields the RNA primer required for the initiation of DNA replication (Figure 1.8) (reviewed by Wagner and Simons, 1994). Inhibition of replication involves the duplex formation between RNAII and its antisense strand RNAI. Binding of RNAI to RNAII results in conformational changes within

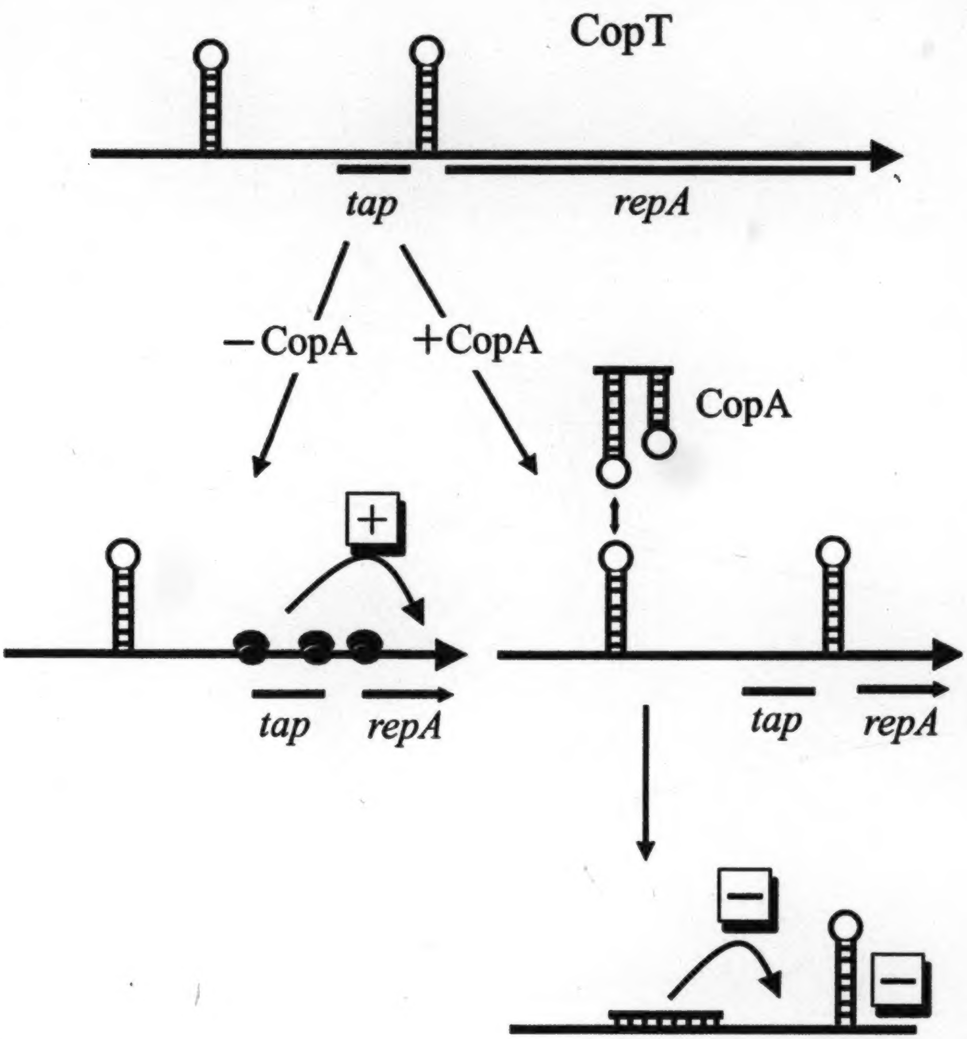


Figure 1.7: CopT/CopA regulatory system of IncFII-like plasmids. CopT transcripts encode for the essential replication factor repA. Translation of RepA is coupled to translation of Tap from an upstream ORF. Association of CopA with CopT occludes the RBS required for Tap translation. In the absence of Tap translation, a stem-loop forms upstream of the repA ORF and occludes the RBS required for RepA translation. Adapted from Wagner & Simons (1994).

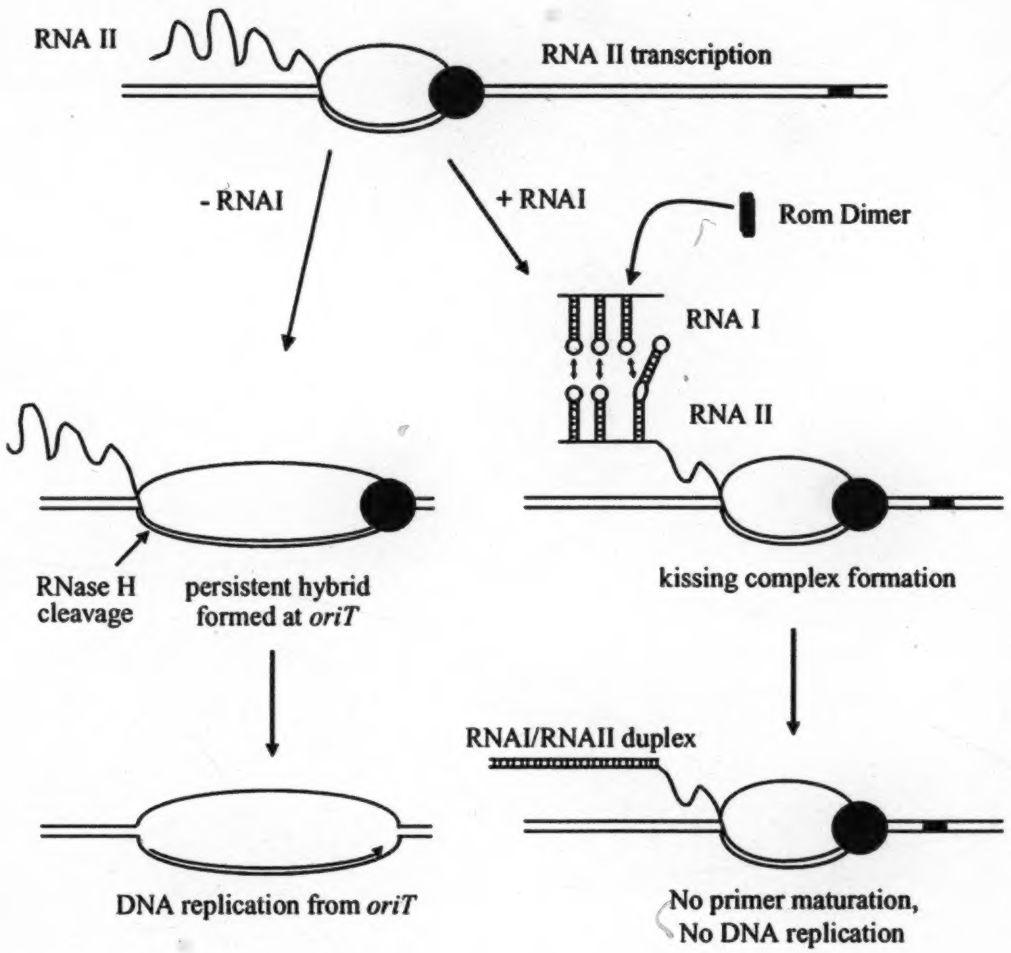

RNA II
RNA II transcription
- RNAI
+ RNAI
Rom Dimer
RNA I
RNA II
RNase H cleavage
persistent hybrid formed at *oriT*
kissing complex formation
RNAI/RNAII duplex
DNA replication from *oriT*
No primer maturation, No DNA replication

Figure 1.8: ColE1 replication regulation by RNAI, RNAII and Rom. RNAII associates with the origin of replication and forms a persistent hybrid. RNase H recognizes the RNA-DNA hybrid and cleaves RNAII to yield a primer for DNA replication. RNAI, assisted by the protein Rom, blocks replication by duplexing with RNAII and preventing RNAII from forming the persistent hybrid at the origin. Adapted from Wagner & Simons (1994).

RNAII that do not allow this RNA to form a persistent hybrid with the DNA and therefore no RNA primer for DNA replication is produced (Masukata & Tomizawa, 1986; Tomizawa *et al.*, 1981; Tomizawa & Itoh, 1981). Duplex formation between RNAI and RNAII initiates through the reversible base pairing between all three loops of RNAI and the corresponding regions on RNAII: loops I and II of RNAI interact with complementary loops I and II of RNAII, while loop III of RNAI interacts with structure IV in RNAII. Kissing between the three loops of RNAI and RNAII keeps the two RNAs together and allows the 5' end of RNAI to start base pairing with its complementary region in the elbow region of RNAII (Tomizawa, 1990). Propagation of base pairing from this 5' end through the rest of the RNA molecules then results in a stable duplex. The ColE1 encoded protein Rom enhances rapid association of RNAI and RNAII and is an essential component of replication inhibition (Eguchi & Tomizawa, 1990, Tomizawa) Tomizawa 1990; Tomizawa & Som, 1984) The structure and function of Rom will be dealt with in more detail in the subsequent section.

1.11 Structure and function of Rom

In the ColE1 RNA antisense regulatory system, Rom acts as a co-regulator by binding to loop-loop interfaces between hairpins of RNAI and RNAII and stabilizing kissing-complex intermediates as well as reducing the rate of dissociation (Eguchi *et al.*, 1991). By binding to the kissing complex, Rom increases the chance of helix propagation through the two complementary strands and increases the rate of duplex formation between RNAI and RNAII. The 1.7 Å resolution structure of Rom revealed that this protein is comprised of two anti-parallel helices, H1 and H2, that homodimerize to form a four-helix bundle (Figure 1.9) (Banner *et al.*, 1987). At present, there are no structural or sequence similarities between Rom and other known RNA binding proteins. Critical residues involved in loop-loop recognition lie in the central region of the helical bundle and are found on the H1-H1' face of the dimer. NMR studies that mapped the interface of Rom in complex with a kissing hairpin (Comolli *et al.*, 1998) between TAR of HIV and its complement TAR*, showed that the central region of the Rom dimer binds to the minor groove formed by the interacting loops, whereas the major groove is occupied by the cluster of phosphates that bridge the major groove and helps to stabilize

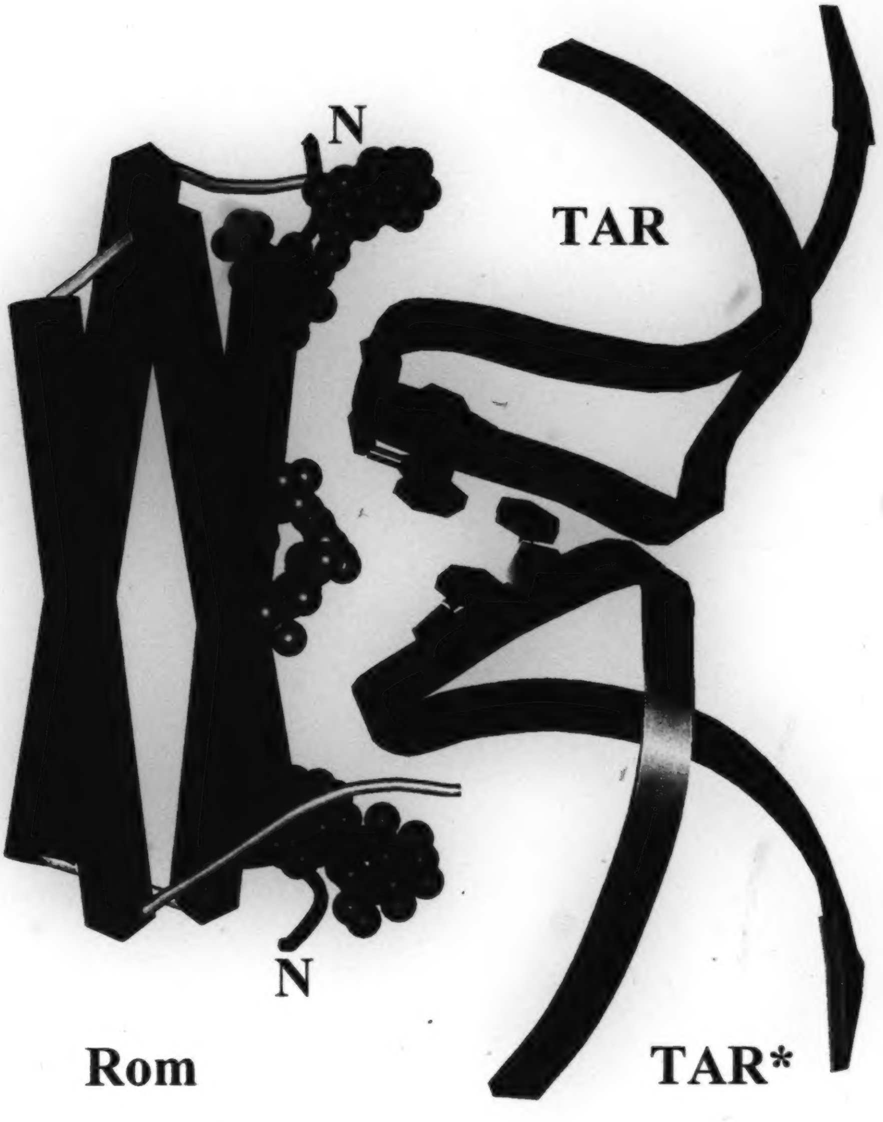


Figure 1.9: Model of Rom bound to a kissing complex. NMR mapping experiments were used to construct a model between Rom bound to the kissing complex formed by TAR and its complement, TAR* (Comolli et al., 1998). The TAR-TAR* kissing complex mimics the kissing complex formed by RNA1 and RNAII. The H1-H1' face of Rom interacts with, and stabilizes, the kissing complex. Phenylalanines (orange) recognize base-pairs (purple) formed between the loops, while lysines (blue) interact with the phophate backbone of the RNA. Adapted from Comolli *et al.* (1998).

the hairpin loops (Figure 1.9). Phenylalanine residues on the H1-H1' face interact with the loop-loop junction and likely stabilize the RNA complex. Furthermore, lysine residues contact the major grooves of the stems on either side of the loop junction. Another interesting aspect of this protein involves the negatively charged patch found opposite to the binding face of Rom and comprised of helices H2 and H2'. This face has been suggested to help promote binding of Rom in the correct orientation with its cognate RNA (Predki *et al.*, 1995).

1.12 Helicases

The double-stranded helix adopted by complementary regions in DNA and RNA is a thermodynamically stable structure that often requires the aid of proteins for strand separation. Unwinding of nucleic acids is observed in numerous biological processes. In the case of DNA, unwinding is required during replication, recombination, DNA repair and RNA transcription (Matson *et al.*, 1994). With the more versatile RNA molecules, unwinding is an important aspect of many RNA-containing biological pathways, where structural rearrangement in RNA is essential for efficient progression of a pathway. Splicing mediated by the spliceosome, for example, requires many strand exchange reactions to occur during the various stages of the pathway (reviewed in Staley & Guthrie, 1998). Other examples include the RNA antisense regulatory systems, such the FinOP system, that require unwinding of intra-molecular stem-loop structures as a prerequisite to the formation of an extended duplex.

Helicases are a major class of enzymes that function to processively and directionally unwind nucleic acids in an ATP-dependent fashion (for recent reviews on helicases see von Hippel & Delagoutte, 2001; Tanner & Linder, 2001). There are seven common sequence motifs that are observed in the various helicases and divergence within these motifs is the basis for defining the families into which the helicases are grouped. Historically, all RNA helicases were categorized into the DEAD-box family, because of the conservation of motif II (DExD/H) within helicases.

Two distinct processes are required for efficient helicase activity: directional translocation of the helicase along the single strand product and unwinding at the junction between single strand and duplex. All helicases require energy from ATP hydrolysis for processive translocation, but unwinding can occur either through an 'active' ATP-dependent mechanism or through the 'passive' ATP-independent thermal fraying of one or two base pairs in front of the translocating helicase (Hippel & Delagoutte, 2001). For example, the PcrA helicase from *Bacillus stearothermophilus* was reported by Soultanas *et al.*, (2000) to use an 'active' mechanism, whereas the HCV helicase was demonstrated to function through a 'passive' mechanism (requiring ATP only for translocation) (Porter *et al.*, 1998).

Another class of proteins that can unwind stretches of duplex in an ATP-independent fashion is the single-stranded nucleic acid binding proteins. These proteins, which include, NCp7 from HIV (described above), the replication protein A (RPA) (Iftode *et al.*, 1999), bacterial SSB (Lohman & Ferrari, 1994), as well as the RNA-binding proteins hnRNP A1 (Herschlag, 1995) and hnRNP C (Shahied *et al.*, 2001), have a high affinity for single stranded nucleic acids and shift the duplex to single strand equilibrium towards the single strand. As a consequence of destabilizing preformed duplexes, single stranded binding proteins can also function as nucleic acid chaperones to facilitate the rate of strand exchange between duplexed regions as has been described above for NCp7.

In this thesis I present evidence that FinO unwinds RNA stem-loop structures through a novel ATP-independent mechanism. Unlike with the single stranded binding proteins, FinO does not recognize single-stranded RNA, but instead appears to function through an induced fit mechanism requiring the initial recognition of double-stranded RNA prior to base-pair separation.

Chapter 2

Experimental Procedures

2.1 Construction of pGEX fusion plasmids

A clone encoding FinO from pR6-5 as a GST fusion (pGEX-FO2) was used as a template for constructing all the various *finO* derivatives (van Biesen & Frost, 1994). PCR products of either the complete *finO* gene or portions of the gene were inserted into a GST-fusion vector (pGEX-KG), between a 5' *Bam*HI site and a 3' *Eco*RI site. To allow direct cloning into pGEX-KG, the upstream and downstream primers used to make the various constructs contained *Bam*HI and *Eco*RI sites, respectively.

Constructs with site-directed mutations were produced using overlap extension as described by Ho *et al.*, (1989). Briefly, pGEX fusions with *finO* or portions of *finO* were used as templates to introduce point mutations. A first round of PCR amplification was used to produce two overlapping products from a given template that together spanned the entire *finO* ORF and were referred to as the 5' and 3' fragments. The 5' fragment is produced using an outside upstream primer (complementary to the 5' end of the template) and an inside downstream primer (complementary to the overlap site). The 3' fragment is produced using an outside downstream primer (complementary to the 3' end of the template) and an inside upstream primer (complementary to the overlap site). The two inside primers are complementary to each other and introduce the mutation of interest. The 5' and 3' fragments from the first PCR round are gel purified and subsequently combined with the two outside primers in a second PCR round to yield an original template that now contains a site specific mutation. The outside primers contained BamHI and EcoRI sites to allow direct cloning back into the pGEX-KG. Additional mutations were introduced sequentially using multiple rounds of overlap extension.

All constructs were sequenced to confirm that we had the desired mutation. Table 2.1 lists all *finO* derived constructs used in these studies and the primers used to produce them.

2.2 Purification of FinO and FinO-derived proteins

FinO and its derivatives were expressed as glutathione-S-transferase (GST) fusions from the pGEX-KG vector (Pharmacia). FinO overexpression plasmids were transformed into *E. coli* DH5α strains. One liter cultures of these strains were grown in LB + 100 µg/mL ampicillin at 37 °C. At an OD_{600} of 0.8, the culture was moved to 25 °C

Table 2.1: List of all primers used in the construction of FinO derivatives

Protein	5' primer(s)[1]	3' primer(s)[1]
native FinO	CCGACGGGATCCATGACAGAGCAGAAGCGACCG	CCGACGGAATTCTTATTTCTCATCAAGCACGGC
FinO(26-186)	CCGACGGGATCCACCATCATCAATGTCACCACG	
FinO(45-186)	CCGACGGGATCCGAGAAGGCTGCCCGGGAAGCA	
FinO(62-186)	CCGACGGGATCCATGCAGGCGCTGTCCATTTAT	
FinO(1-61)		CCGACGGAATTCTTATCTGGCCTGCGCTTTTTT
FinO(1-174))		CCGACGGAATTCTTACCGGTTCTGGCGGCGGAT
FinO(62-174)	CCGACGGGATCCATGCAGGCGCTGTCCATTTAT	CCGACGGAATTCTTACCGGTTCTGGCGGCGGAT
FinO(62-170)	CCGACGGGATCCATGCAGGCGCTGTCCATTTAT	CCGACGGAATTCTTAGCGGATTTTATCCAGACG
FinO(26-186)L124M	CCGACGGGATCCACCATCATCAATGTCACCACG	
	GCGCAGGGCGATGAAGGCCATC	GATGGCCTTCATCGCCCTGCGC
FinO(26-186)L96,124M	CCGACGGGATCCACCATCATCAATGTCACCACG	
	GCGCAGGGCGATGAAGGCCATC	GATGGCCTTCATCGCCCTGCGC
	GACACGCCCCGGCTTATGGCCTGCGGT	ACCGCAGGCCATAAGCCGGGGCGTGTC
FinO(3C→S)[2]	GCTTCTGGCCTCCGGTATCCGGG	CCCGGATACCGGAGGCCAGAAGC
	AGCTATCTGTCTGCCATGAAAGCC	GGCTTTCATGGCAGACAGATAGCT
	AGCCGGTGCCTCCCGGTATGACA	TGTCATACCGGGAGGCACCGGCT
FinO(3C→S)K37C	GCCACCAAAATGGTGTGTGAAAAAGCAGA	TCTGCTTTTTCACACACCATTTTGGTGGC
FinO(3C→S)K40C	ATGGAAGGTGAAATGTCAGAAACTGGCGG	CCGCCAGTTTCTGACATTTCACCTTCCAT
FinO(3C→S)K42C	GGTGAAAAAGCAGTGTCTGGCGGAGAAGG	CCTTCTCCGCCAGACACTGCTTTTTCACC
FinO(3C→S)K46C	GAAACTGGCGGAGTGTGCTGCCCGGGAAG	CTTCCCGGGCAGCACACTCCGCCAGTTTC
FinO(3C→S)R81C	TGAACACCCTGTGTCCCTGGTGGCC	GGCCACCAGGGACACAGGGTGTTCA
FinO(3C→S)R118C	CGCTCTCGCATTGTAAACTGCGCAGG	CCTGCGCAGTTTACAATGCGAGAGCG
FinO(3C→S)K121C	CGCATAAAAAACTGTGTAGGGCGCTGAA	TTCAGCGCCCTACACAGTTTTTTATGCG
FinO(3C→S)K125C	CAGGGCGCTGTGTGCCATCACCCG	CGGGTGATGGCACACAGCGCCCTG
FinO(3C→S)K135C	AAGCTATCTGTGTGCCATGAAAGCC	GGCTTTCATGGCACACAGATAGCTT
FinO(3C→S)K142C	AGCCGGTGCCTGCCGGTATGACA	TGTCATACCGGCAGGCACCGGCT
FinO(3C→S)K147C	GGTATGACACGTGTGGGTATGTGACG	CGTCACATACCCACACGTGTCATACC

continued on next page ....

Table 2.1: continued

Protein	5' primer(s)[1]	3' primer(s)[1]
FinO(3C→S)K165C	TATGCGGCAGAGTGTCTGGATAAAATC	GATTTTATCCAGACACTCTGCCGCATA
FinO(3C→S)K170C	CTGGATAAAATCTGTCGCCAGAACCGG	CCGGTTCTGGCGACAGATTTTATCCAG
FinO(3C→S)K176C	CAGAACCGGATATGTGCAGAACTTCAG	CTGAAGTTCTGCACATATCCGGTTCTG
FinO T26A,I27A[3]	CGGAGCCCGGAAAGCCGCCATCAATGTCACC	GGTGACATTGATGGCGGCTTTCCGGCTCCG
FinO V30A,T31A	ACCATCATCAATGCCGCCACGCCACCAAAA	TTTTGGTGGCGTGGCGGCATTGATGATGGT
FinO T32A/P33A	ATCAATGTCACCGCGGCACCAAAATGGAAG	CTTCCATTTTGGTGCCGCGGTGACATTGAT
FinO P34A,K35A	GTCACCACGCCAGCAGCATGGAAGGTGAAA	TTTCACCTTCCATGCTGCTGGCGTGGTGAC
FinO W36A	ACGCCACCAAAAGCGAAGGTGAAAAAG	CTTTTTCACCTTCGCTTTTGGTGGCGT
FinO K37A,V38A	CCACCAAAATGGGCGGCCAAAAAGCAGAAA	TTTCTGCTTTTTGGCCGCCCATTTTGGTGG
FinO K39A,K40A	AAATGGAAGGTGGCAGCGCAGAAACTGGCG	CGCCAGTTTCTGCGCTGCCACCTTCCAATTT
FinO Q41A,K42A	AAGGTGAAAAAGGCGGCACTGGCGGAGAAG	CTTCTCCGCCAGTGCCGCCTTTTTCACCTT
FinO L43A	AAAAAAGCAGAAAGCGGCGGAGAAGGCTGCC	GGCAGCTTCTCCGCCGCTTTCTGCTTTTT
FinO E45A, L46A	CAGAAACTGGCGGCGGCGGCTGCCCGGGAA	TTCCCGGGCAGCCGCCGCCGCCAGTTTCTG
FinO T32A	ATCAATGTCACCGCGCCACCAAAATGG	CCATTTTGGTGGCGCGGTGACATTGAT
FinO P33A	AATGTCACCACGGCACCAAAATGGAAG	CTTCCATTTTGGTGCCGTGGTGACATT
FinO P34A	GTCACCACGCCAGCAAAATGGAAGGTG	CACCTTCCATTTTGCTGGCGTGGTGAC
FinO K35A	ACCACGCCACCAGCATGGAAGGTGAAA	TTTCACCTTCCATGCTGGTGGCGTGGT
FinO K37A	CCACCAAAATGGGCGGTCAAAAAGCAG	CTGCTTTTTGACCGCCCATTTTGGTGG
FinO V38A	CCAAAATGGAAGGCCAAAAAGCAGAAA	TTTCTGCTTTTTGGCCTTCCATTTTGG
FinO K39A	AAATGGAAGGTGGCAAAGCAGAAACTG	CAGTTTCTGCTTTGCCACCTTCCATTT
FinO K40A	TGGAAGGTGAAAGCGCAGAAACTGGCG	CGCCAGTTTCTGCGCTTTCACCTTCCA
FinO Q41A	AAGGTGAAAAAGGCGAAACTGGCGGAG	CTCCGCCAGTTTCGCCTTTTTCACCTT
FinO K42A	GTGAAAAAGCAGGCACTGGCGGAGAAG	CTTCTCCGCCAGTGCCTGCTTTTTCAC

---

[1] Unless otherwise specified, the outisde 5' and 3' primers are the same ones used for cloning native FinO (from plasmid R6-5).

All primers are shown with the 5' end on the right. Mutation sites on inside primers are indicated in bold lettering.

Restriction sites on outside primers are underlined

[2] All FinO(3→S) clones with additional mutations were constructed using the FinO(3C→S) clone as a template

[3] FinO constructs containing double and single alanine point mutant were created by Ross Edwards and David Arthur

and induced with 0.2 mM IPTG. After 5 h of growth, cells were harvested and stored at –70 °C overnight. Cells were resuspended in 50 mL of 50 mM potassium phosphate (pH 6.5), 150 mM NaCl, 1 mM EDTA, 0.1% βME, 0.1 mg/mL PMSF, 1 μg/mL pepstatin, 1 μg/mL leupeptin, and 0.1 mg/mL lysozyme, and the cell solution was mixed for 30 min, at 4 °C, to break down the cell wall. Lysis of cells was completed by brief sonication. The lysate was spun at 30000*g* to remove insoluble debris, and the cleared lysate was loaded onto a 10 mL glutathione-agarose column (Pharmacia). The column was subsequently washed with a solution of 50 mM Tris-HCl (pH 8.5), 150 mM NaCl, and 1 mM EDTA, and GST-FinO was eluted with the same buffer containing 20 mM reduced glutathione. Fractions containing FinO were pooled (about 25 mL) and were incubated with 17 units of thrombin for ~16 h to cleave GST from FinO. The digestion was stopped by addition of PMSF to a final concentration of 10 μg/mL. FinO was purified from GST by cation exchange chromatography using a 30 mL Fast-SP column (Pharmacia) and a 10 column volume gradient from 0 to 1 M NaCl in a buffer containing 50 mM potassium phosphate (pH 6.5), 1 mM EDTA, and 0.1% βME. FinO fragments eluted from this column at approximately 400 mM NaCl. Fractions containing purified FinO (determined by polyacrylamide gel electrophoresis) were diluted 3-fold with the 50 mM potassium phosphate (pH 6.5) buffer and were bound to a 1 mL Fast-SP column. Concentrated FinO was eluted from this column with 1-2 mL of 600 mM NaCl, 50 mM potassium phosphate (pH 6.5), 1 mM EDTA, and 0.1% βME. Protein concentrations were determined by the Bradford assay (BIORAD), which was calibrated for the true molar concentration of each of the protein fragments by amino acid analysis.

Selenomethionine-substituted FinO(26–186)L96,124M was expressed by incorporation of selenomethionine into the bacterial growth media, under growth conditions that repress methionine biosynthesis (Doublie, 1997). Induction of the culture and the subsequent purification of selenomethionine-substituted protein was performed as described above for FinO and FinO derivatives.

Experiments presented in chapter 6 used proteins that upon purification, were divided into 50 μL aliquots at 3 mg/mL, or less, flash frozen in liquid nitrogen and stored at –70 °C until required.

2.3 Preparation of RNA substrates

In vitro transcription of FinP and *traJ* mRNA from linearized plasmid templates has been described elsewhere (Jerome *et al.*, 1999). The following is the procedure used to produce all other RNA substrates used in these studies. RNA was transcribed *in vitro* using T7 RNA polymerase and a synthesized DNA template containing a T7 promoter site and a top strand annealed to this promoter region. Sequences of the templates and top strand are given in table 2.2, while the sequences of the transcribed RNA substrates are shown in figure 6.1. The transcription reaction was performed at 37 °C for two hours in 40 mM Tris (pH 8.1), 1 mM spermidine, 0.01% Triton X-100, 5 mM DDT, 80 mg/ml PEG 8K, 4.3 mM of ATP, CTP, GTP and UTP, and 11 mM $MgCl_2$ and containing 0.4 ug/ml pyrophosphatase (Sigma), 12.5 ug/ml of T7 RNA polymerase (courtesy of Mark Glover) and 300 nM of the template annealed to the top strand. The transcription reaction was prepared at room temperature. This avoids precipitation of nucleic acids by spermidine at lower temperatures. Following the transcription reaction, 400 mM of sodium acetate was added to the mixture and RNA was ethanol precipitated. The pellet formed was washed in 70% ethanol, resuspended with gel loading buffer (7M urea and 0.5x TBE) and heated at 85 °C for 5 minutes. RNA was then subjected to denaturing 20%-PAGE for approximately four hours. The band corresponding to the transcribed RNA was observed by UV-shadowing and excised from the gel. RNA was then eluted from the gel using the EluTrap from (Schleicher & Schuell), concentrated to approximately 1 mg/ml using a 5 kDa molecular weight cutoff Ultrafree centrifugal filter device (Millipore) and stored at –70 °C until required.

SII and SII+4 duplexes were prepared as follows: 5 nM of the ^{32}P-labeled strands were combined with an excess of their complementary cold strand (50 nM) in 10 mM Tris-HCl (pH 8.1) and 100 mM NaCl and annealed by slow cooling to 23 °C from an initial temperature of either 85 °C for SII, 90 °C for SII+4, or 100 °C for $SLII_x/SLIIc_x$. For the strands in the RNA duplexes, their corresponding DNA templates were designed to allow for maximal transcriptional yields, while maintaining base-pair complementarity in the RNA duplex.

The fluorescein-labeled duplexes were equivalent to the SII duplexes, except for the presence of a 5' fluorescein on either one or the other strand in the duplex. The

Table 2.2: List of DNA templates used in synthesis of RNA substrates

RNA	DNA template[1]
SLII(A)	GGTCCTGCATCGACTGTC**CTATAGTGAGTCGTATTA**
SLII(B)	AAAATCGCCGATGCAGGAC**CTATAGTGAGTCGTATTA**
SLII+4(A)	GGAACTCCCTGCATCGGCTGTC**CTATAGTGAGTCGTATT**
SLII+4(B)	AAAATCGCCGATGCAGGGAGTTC**CTATAGTGAGTCGTATTA**
SLII	AAAATCGCCGATGCAGGGAGACGTGAACTCCCTGCATCGACTGTC**CTATAGTGAGTCGTATTA**
$SLII_x$	GGACTCGCCGATGCAGGGAGACGTGAACTCCCTGCATCGACTGTC**CTATAGTGAGTCGTATTA**
$SLIIc_x$	GGACAGTCGATGCAGGGAGTTCACGTCTCCCTGCATCGGCGAGTC**CTATAGTGAGTCGTATTA**
top strand	TAATACGACTCACTATAG

---

[1] region annealed to top strand is indicated in bold type. All primers are with the 5' end on the right.

fluorescein-labeled RNA was purchased from Dharmacon Research Inc., while the non-labeled strands were produced by *in vitro* transcription, as described above. Mass spectroscopic analysis of the fluorescein labeled strands indicated that >90 % of the RNAs contained fluorescein. Strands were annealed by slow cooling from an initial temperature of 85 °C to room temperature over a two-hour period. Annealing was performed in the dark (to avoid bleaching of fluorescein) with approximately 45 μM of the two complementary strands in 1 mM EDTA, 10 mM Tris (pH 8.0) and 100 mM KCl. RNA duplexes were stored at –20 °C. Individual aliquots of duplex were only used once to avoid freeze-thaw and degradation.

2.4 Proteolysis

Reactions contained 200 μg (9 nmol) of FinO, 230 mM NaCl, 19 mM potassium phosphate (pH 6.5), 11 mM Tris-HCl (pH 8.1), 38% glycerol, 1 mM EDTA, 0.1% βME, and, in some experiments, 410 μg (28 nmol) of SLII RNA, in a total volume of 105 μL. Samples were equilibrated at 25 °C for 30 min. For 25 °C reactions, 4 μL of a 50 μg/mL stock of trypsin was added to initiate the reactions. For 4 °C reactions, samples were incubated at 4 °C for a further 10 min, and 4 μL of 1 mg/mL trypsin was subsequently added. Five microliter aliquots were taken at various times, and 1 μL of 1 mg/mL PMSF was added to stop the digestion. Samples were analyzed by 15% SDS-PAGE. To identify the various proteolytic fragments, trypsin digests were scaled up to 250 μL. At the desired time point, the reactions were stopped with 3 μL of 1 mg/mL PMSF. Five microliters of this reaction was analyzed by SDS-PAGE. The molecular masses of the proteolytic fragments in the remainder of the reaction were determined by electrospray mass spectroscopy using a VG Quattro triple quadrupole mass spectrometer. The proteolytic fragments FinO(62-170) and FinO(62-174) (Figure 3.3) were first purified prior to mass spectroscopy by HPLC using a Zorbax C8 reversed-phase column. The first four N-terminal amino acids of these fragments were determined using an HP G1005A amino acid sequencing system with routine 3.0 chemistry and biphasic column technology at the Alberta Peptide Institute.

2.5 Circular dichroism spectroscopy

CD spectroscopy for FinO and FinO-derived fragments was performed at a concentration of approximately 1 mg/mL in a buffer containing 50 mM MES (pH 6.5), 150 mM NaCl, 5 mM EDTA, and 1% βME. Exact protein concentrations were determined by amino acid analysis. CD spectra were determined either in the far-UV (190-255 nm) or in the near-UV (255-320 nm) ranges. Protein solutions were loaded into calibrated 0.02 and 1 cm fused silica cells for the far- and near-UV analysis, respectively. Spectra were recorded on a Jasco J-720 spectropolarimeter (Jasco Inc., Easton, MD) interfaced with an Epson Equity 386/25 computer and controlled by Jasco software. A Lauda water bath (Brinkmann Instruments) was used to control the cell temperature. The spectropolarimeter is routinely calibrated with ammonium d-(+)-10-camphor-sulfonate at 290.5 and 192 nm. For the far-UV data, mean molar ellipticity is calculated as the ellipticity per mole of protein, rather than per mole of amino acid residue, to facilitate comparison of the spectra obtained from different FinO fragments. The near-UV data are presented as the mean molar ellipticity per mole of aromatic residues. Each data point is the average of 10 samplings. The helical content for each fragment was estimated by Provencher-Glöckner analysis (Provencher & Glöckner, 1981).

2.6 Electrophoretic mobility shift assay (EMSA)

For experiments performed with full-length FinP RNA, 7.5 fmol of ^{32}P-labeled FinP was incubated with increasing concentrations of FinO deletion proteins in separate reactions. RNA and protein were mixed in a total volume of 30 µL containing 50 mM Tris-HCl (pH 8), 1 mM EDTA, 100 mM NaCl, 3.0 µg of RNase-free BSA (Pharmacia), 2 mM DTT, 10% glycerol, and 7.6 units of RNAguard (Pharmacia). Reactions were incubated for 30 min at 4 °C. In competition assays, *E. coli* total tRNA (1000-fold molar excess vs FinP) was added to reactions containing FinP RNA and incubated at 4°C for 5 min before addition of the protein. Reactions were loaded onto a continuously running 5% or 8% nondenaturing polyacrylamide gel and electrophoresed for 1 h at 150V, 4°C. Polyacrylamide gels contained Tris-glycine buffer (25 mM Tris, 0.19 M glycine, pH 8-8.3), which was also used as the running buffer. Gels were imaged on Molecular

Dynamics storage phosphor screens with a Molecular Dynamics Phosphorimager 445 SI and quantitated using Molecular Dynamics ImageQuaNT software.

The EMSAs described in chapter 3, using SLII were performed in a similar manner as described for FinP RNA. Each reaction contained 40 μM ^{32}P-SLII RNA and 2.5 μM GST-FinO and/ or FinO. Proteins were mixed first and incubated for 30 min at 25 °C followed by another 30 min incubation at 4 °C. SLII was then added, and the samples were incubated for another 30 min at 4 °C. Gel electrophoresis was performed as described above, and the positions of RNA-containing species within the gel were visualized by autoradiography.

In chapter 5, the EMSAs performed using SLII or SII are similar to those described above for FinP, with the following modification; binding reactions contained 5 μL of binding buffer (50 mM Tris (pH 8.1), 100 μg/mL BSA, 10% glycerol, 0.1 % β-mercaptoethanol and 40 mM NaCl), 4 μL of protein in 50 mM MES (pH 6.5), 0.1% BME, 450 mM NaCl and 100 μg/ml BSA, and 1 μL of labeled RNA at a concentration of 500 pM.

The apparent equilibrium association constant (K_a) described in the text for FinO and FinO derived proteins in complex with RNA were calculated from the equilibrium expression:

$$K_a = \frac{[\mathrm{FinO} \bullet \mathrm{FinP}]}{[\mathrm{FinO}][\mathrm{FinP}]}$$

where [FinO•FinP], [FinO], and [FinP] refer to the equilibrium concentrations of FinO•FinP complex, free FinO, and free FinP, respectively. The ratio of FinO-bound FinP to free FinP for each binding reaction was determined from the ratio of the intensities of the bands corresponding to the bound and free FinP species. Because FinO was in vast molar excess over FinP in all reactions, the total FinO concentration was assumed to be equivalent to [FinO]. K_a values were determined for each binding reaction and averaged for a given titration experiment. Unless otherwise specified, all K_a determinations were obtained from at least three separate titration experiments.

2.7 Crystallization of FinO(26-186)

All crystals were grown at 4 °C by the hanging drop, vapour diffusion method. For the hanging drop, 1 μL of a 10 mg mL^{-1} protein solution in 150 mM NaCl, 50 mM 2-(4-morpholino)-ethane sulfonic acid (pH 6.5), 0.1% βME (v/v) and 1 mM EDTA was mixed with 1 μL of reservoir solution (12% PEG 4000 and 50 mM Tris, pH 7.2). Crystals of FinO(26–186) grew as thin plates to a maximum size of 400 μm x 200 μm x ~5 μm within seven hours. Crystals of FinO(26–186)L124M and selenomethionine substituted FinO(26–186)L96,124M were nucleated with reservoir solution containing crushed FinO(26–186) crystals. The maximum size of these seeded crystals was approximately the same as for FinO(26–186).

2.8 Crystallographic data collection and processing

Crystals were harvested in reservoir solution and transferred to a series of reservoir solutions with increasing glycerol concentrations to a maximum concentration of 25% (v/v). These crystals were then flash frozen directly in a 100 K N_2 stream for data collection. The MAD data set was collected with a Brandeis 1K CCD detector at beamline X12C at the NSLS. Data from FinO(26–186)L124M crystals were collected at beamline 14-BM-D at the Advanced Photon Source (APS) using a ADSC Quantum-4 CCD detector. All data were indexed and scaled using Denzo and Scalepack (Otwinowski & Minor, 1997). The selenomethionine positions and initial MAD phases were calculated using the program SOLVE, treating the data as a special case of multiple isomorphous replacement with anomalous scattering, and using the inflection wavelength data as a pseudo-native data set (Terwilliger, 1987). Phases were further improved by solvent flattening and histogram matching as implemented in the program DM (Cowtan, 1994). The solvent mask was edited to include the solvent exposed helix 1.

2.9 Model building and refinement of FinO(26-186) structure

The MAD phased, solvent flattened electron density map calculated to 2.6 Å resolution gave excellent electron density for the region corresponding to residues 54–184 and poorer but interpretable density for the region corresponding to residues 38–53. Using this map, we built a model incorporating most of the side chains for residues 38–

183, using the programs O (Jones *et al.*, 1991) and Xtalview/Xfit (Jones *et al.*, 1991; McRee, 1999). In the early stages of refinement we used the program X-PLOR (Brünger, 1992) for restrained individual B-factor refinement, positional refinement and simulated annealing protocols. After three rounds of refinement, the R factor was brought down to 32.7 % (R_{free} 37.3%). Residues 34 to 37 and 184 were now introduced into our model and refinement was continued using the maximum likelihood torsion angle dynamics algorithm from the CNS package, employing a solvent mask and an overall anisotropic B-factor correction (Brünger, 1999). At an R-factor of 30% (R_{free} 34.7%), additional refinement was continued with the 2.0 Å data set collected from the FinO(26–186)L124M crystals. The final working R-factor was 19.7% (R_{free} 22.4%). Figures were prepared with the programs BOBSCRIPT (Esnouf, 1997), MOLSCRIPT (Kraulis, 1991), RASTER3D (Esnouf, 1997; Merritt, 1997) and GRASP (Nicholls *et al.*, 1991).

The atomic coordinates of the final refined FinO(26-186) model have been deposited with the Protein Data Bank (accession code 1DVO).

2.10 RNA duplexing assays

Duplex analyses between FinP and *traJ*-mRNA were performed as described previously (Sandercock & Frost, 1998) with minor changes. Briefly, 0.15 nM ^{32}P-FinP and ^{3}H-*tra*J184 RNA, in 10-fold molar excess, were incubated in 50 µL of duplexing buffer (25 mM Tris-HCl, pH 8.0, 4 mM magnesium acetate, 0.4 mM EDTA, 40 mM NaCl, 40 µg mL^{-1} RNase-free bovine serum albumin, 0.3 units mL^{-1} RNasin) in the presence or absence of various FinO deletion derivatives. In some cases, a two-fold or five-fold molar excess of *tra*J184 RNA was used in order to produce measurable rates of duplex formation. Aliquots (5 µL) were removed at specified time points and added to 10 µL ice-cold stop solution (95% formamide, 10 mM EDTA, 0.5% each xylene cyanol and bromophenol blue) and electrophoresed on 8% native polyacrylamide gels.

Duplexing between $SLII_x$ and $SLIIc_x$ was performed in 50 µl reaction mixtures containing 25 µl of 2x reaction buffer (50 mM Tris (pH 8.1), 100 µg/mL BSA, 10% glycerol, 0.1 % βME and 40 mM NaCl), 5 µl of protein at 10 µM, in 20mM MES (pH 6.5), 0.1% βME, and 60 mM NaCl, 5 µl of $SLIIc_x$ at 1 µM, 5 µl of ^{32}P-labeled $SLII_x$ at 50 nM and 10 µl of ddH_2O. To initiate duplexing, labeled $SLII_x$, preincubated to 37 °C, was added to reaction mixtures also preincubated at 37 °C and lacking SLII_x. 5 µL aliquots were taken at various time points and added to 5 µL of cold stop buffer (5% glycerol, 0.4% SDS and 20 mM EDTA). Samples were subjected to 10% non-denaturing PAGE for 2 hours and the resulting bands, corresponding to duplexed and free ^{32}P-labeled, were visualized by exposing a Molecular Dynamics storage phosphor screen to the gels for 20 hours and then scanning the screen using the Storm 840 phosphorimager (Molecular Dynamics). Bands were quantified with ImageQuanNT software (Molecular Dynamics).

The second order apparent rate constant (k_2) for duplex formation was determined essentially as described (Persson *et al.*, 1988). Briefly, the second order rate equation for the duplexing reaction is,

$$\frac{d[A*B]}{dt} = -\frac{d[A*]}{dt} = k_2[A*][B]$$

where [A*B], [A*], and [B] are the concentrations of duplex, ^{32}P-labeled strand A and unlabeled strand B, respectively, at time t. When the unlabeled strand is in large excess over the labeled strand ($[B]_o >> [A*]_o$) , duplexing can be expressed as a pseudo-first order reaction,

$$-\frac{d[A*]}{dt} = k_1[A*]$$

and k_1 can be determined from a plot of ln[A*] vs t. k_2 is then determined from k_1 by the equation $k_2 \cong k_1/[B]_o$.

2.11 Unwinding assay

Unwinding assays were set-up on ice in 10 µL reaction volumes consisting of 5 µL of the reaction buffer used in the duplexing assays, (50 mM Tris (pH 8.1), 100 µg/mL BSA, 10% glycerol, 0.1 % β-mercaptoethanol and 40 mM NaCl), 1.5 µL protein in 20 mM MES (pH 6.5), 0.04% β-mercaptoethanol and 60 mM NaCl, 1 µL of the labeled SII duplex (5 nM duplex and 45 nM of the excess cold down strand), 1 µL of the unlabeled strand (1.25 µM) and 1.5 µL of distilled water. Unwinding assays were initiated by placing reaction tubes at 23 °C, 30 °C or 37 °C, as indicated. Reactions were stopped by addition of an equal volume of stop solution (5% glycerol, 0.4% SDS and 20 mM EDTA) to aliquots taken from the reaction mixture at various time points for the rate determination assays, or at the two hour time point for all the other unwinding experiments. Samples were subjected to 15% non-denaturing PAGE at room temperature to separate the free and duplexed labeled (A) strand. For rate determination assays, samples were loaded onto a continuously running gel. To visualize bands corresponding to duplexed and free ^{32}P-labeled, a Molecular Dynamics storage phosphor screen was exposed to the gels for 20 hours and scanned using the Storm 840 phosphorimager (Molecular Dynamics). Bands were quantified with ImageQuanNT software (Molecular Dynamics). Unwinding rates were determined by fitting data to the equation,

$$[A^*] = [A^*B]_o(1 - e^{kt}) + c$$

where k is the apparent first order rate constant of unwinding, c is the constant of integration, $[A^*B]_o$ is the initial concentration of duplex RNA, and [A*] is the concentration of ^{32}P-labeled single stranded RNA relative to $[A^*B]_o$ (calculated as the ratio of band intensities of the single stranded RNA divided by the duplex RNA, and then the ratio at each reaction time (t) was normalized to the zero time point, which was set to zero).

2.12 Protein-RNA cross-linking

The cross-linker APA-Br (Sigma) was initially dissolved in methanol to a final concentration of 208 mM. To attach APA to the cysteine mutants, 1 µL of the APA stock was added to 100 µL of an 80 µM protein solution containing the buffer 10 mM Tris (pH 7.0), 600 mM NaCl and 1 mM EDTA. The reaction mixture was then incubated in the dark for two hours at room temperature. Excess APA was subsequently removed using a BIORAD P-30 spin column pre-equilibrated with 10 mM Tris (pH 7.0), 600 mM NaCl and 1mM EDTA. We used a dithionitrobenzoate (DTNB, Sigma) assay (described in Hall & Fox, 1999) to determine the reactivity of the cysteine residues at pH 7.0. This assay revealed that at least 90% of the thiol groups were accessible to DTNB for each cysteine mutant.

The cross-linking reactions were performed with 42 μM protein and 81 μM SLII that had been pre-incubated for 10 min. at 4 °C, in 10 mM Tris (pH 7.0), 600 mM NaCl and 1 mM EDTA. Reactions were performed in a 96-well plate that was placed on ice under a 302 nm UV-light source (115V, 60Hz and 160 mA) at a distance of approximately 4 cm. Samples were exposed to UV-light for 10 minutes, mixed with 5x load buffer (250mM Tris pH 6.8, 40% glycerol, 8% sodium dodecyl sulfate, 0.4% bromophenol blue) and electrophoresed for one hour, at 150 V on a 15% denaturing polyacrylamide gel. Gels were then stained with ethidium bromide to allow detection of the RNA, followed by Coomassie staining to visualize protein.

2.13 GelFRET assay

Texas Red C5 bromoacetamide (TR, Molecular Probes) was dissolved in DMSO to a final concentration of 20 mM, divided into 10 μL aliquots and stored at –20 °C in the dark. For modification of the cysteine point mutants, 100 μL of protein at a concentration of 10 μM in 50 mM Tris (pH 8.0), 600 mM NaCl and 1 mM EDTA, was mixed with 1 μL of 20 mM TR and incubated at room temperature for two hours. The reaction mixture was subsequently separated over a BioRad P30 spin-column to remove unincorporated TR. Some free TR passed through the spin column with the protein, with the amount getting through varying from sample to sample. This did not affect the gelFRET results, since free TR had a different electrophoretic mobility than the nucleoprotein complexes (Figure 5.5). FinO-RNA complexes were formed by the incubation of approximately 500 nM of protein with 370 nM of duplex, in 50 mM potassium phosphate (pH 7), 450 mM NaCl and 15% sucrose, at 4 °C for at least 10 minutes. The complexes were then separated from free components by native 10% PAGE at 4 °C and 250 V for 3 hours. The gels were analyzed using a FluorImagerFSI fluorescence scanner (Molecular Dynamics), with a 488 nm argon ion laser to excite the fluorescein on the RNA, a 530 ± 15 nm band pass filter to detect fluorescein fluorescence, and a 610 nm long-pass filter to detect TR fluorescence. Calibration standards containing only donor or acceptor fluorophores were used to determine the ratio of fluorescence emissions for each fluorophore through each filter.

To determine the percentage of Texas Red modified protein, TR reacted proteins were first separated from free TR by 15% SDS-PAGE, and visualized by UV excitation of Texas Red or Coomassie staining. The intensity of the protein bands, upon UV excitation, were normalized to the intensity of Coomassie stained bands, assuming that the R165C mutant, which displayed the highest ratio of UV excitation intensity to Coomassie intensity of all the mutants, was 100% modified. We estimate that all mutants were at least 60% modified.

2.14 Mating assays

Mating assays were performed essentially as described by Sandercock & Frost (1998) using *Escherichia coli* MC4100 cells bearing the F derivative plasmid pOX38-Km and various pGEX-FinO plasmids. The presence of both plasmids was confirmed by agarose gel electrophoresis of plasmid DNA isolated from the *E. coli* strains, and GST-FinO protein expression levels were assayed by Western blot analysis using anti-GST antibodies (Sigma) and anti-FinO antiserum. All GST-FinO proteins used in these studies were expressed at similar levels, within +/- 20% of wild type. The ratio of transconjugants to donors was calculated, allowing mating efficiency to be compared with the control of conjugal transfer of pOX38-Km alone.

2.15 Northern blot analysis to determine FinP half-life

The half-life of FinP RNA isolated from the GST-FinO-expressing *E. coli* MC4100 strains used for the mating assays was assessed by Northern blot analysis as previously described (Sandercock & Frost, 1998; Jerome *et al.*, 1999). An equivalent amount (35 μg) of total RNA was loaded in each lane of the gel used for the Northern analysis. For FinP half-life determinations, FinP band intensities were measured and were normalized based on the amounts of tRNA^{ser} detected in same lane

Chapter 3

The FinO Repressor of Bacterial Conjugation Contains Two RNA Binding Regions*

* A version of this chapter is published in; *Biochemistry*, Vol. 38, No. 42, 1999, Alexandru F. Ghetu, Michael J. Gubbins, Kimio Oikawa, Cyril M. Kay, Laura S. Frost, and J. N. Mark Glover.

3.1 Introduction

The F-plasmid of *Escherichia coli* is transferred to recipient cells via the gene products encoded in the approximately 35 kb transfer *(tra)* region of the plasmid (reviewed in Frost *et al.*, 1994). The TraJ protein encoded within this region is known to be a positive regulator of transcription of the F *tra* genes from the P_Y promoter located at the start of the transfer region (Cuozzo & Silverman, 1986). The FinOP (fertility inhibition) system of F and F-like plasmids controls the ability and frequency with which these plasmids are transferred from donor to recipient *E. coli* (Finnegan & Willetts, 1972). Control of TraJ expression, and therefore of *tra* gene expression and of F-plasmid transfer, is mediated by two components comprising the FinOP system, FinP and FinO (Finnegan & Willetts, 1972). FinP is an approximately 79 nucleotide RNA molecule that is complementary to the untranslated leader of the *traJ* mRNA (Mullineaux & Willetts, 1985). This antisense RNA molecule contains two stem-loops which are complementary to the analogous stem-loops in *traJ* mRNA (Figure 3.1B). FinP is thought to form an RNA/RNA duplex with *traJ* mRNA, occluding the *traJ* ribosome binding site located in the 5' stem-loop of the mRNA and preventing translation of the TraJ protein (van Biesen & Frost, 1994).

The FinO protein encoded by a variety of F-like plasmids (Cram *et al.*, 1991; McIntyre & Dempsey, 1987; van Biesen & Frost, 1992) is a basic, 186-residue protein with a molecular mass of approximately 21.2 kDa (Figure 3.1A) (Yoshioka *et al.*, 1987; Yoshioka *et al.*, 1990; Sandercock & Frost, 1998). In F, the FinO gene is interrupted by an IS3 insertion. This causes F to be derepressed for transfer (Yoshioka *et al.*, 1987; Cheah & Skurray, 1986) and demonstrates the requirement for FinO in promoting the repression of F transfer. FinO stabilizes FinP in vivo, presumably by preventing its degradation by RNase E, thus increasing its *in vivo* concentration (Lee *et al.*, 1992; Jerome *et al.*, 1999). FinO can then bind to both FinP and *traJ* RNA, promoting duplex formation and blocking TraJ translation (Sandercock & Frost, 1998; van Biesen *et al.*, 1993).

Recent biochemical studies have shed light on how FinO selectively binds to both TraJ and FinP RNAs (Jerome & Frost, 1999). The lack of sequence similarity between these RNAs suggests that FinO does not bind in a sequence-specific manner. Instead,

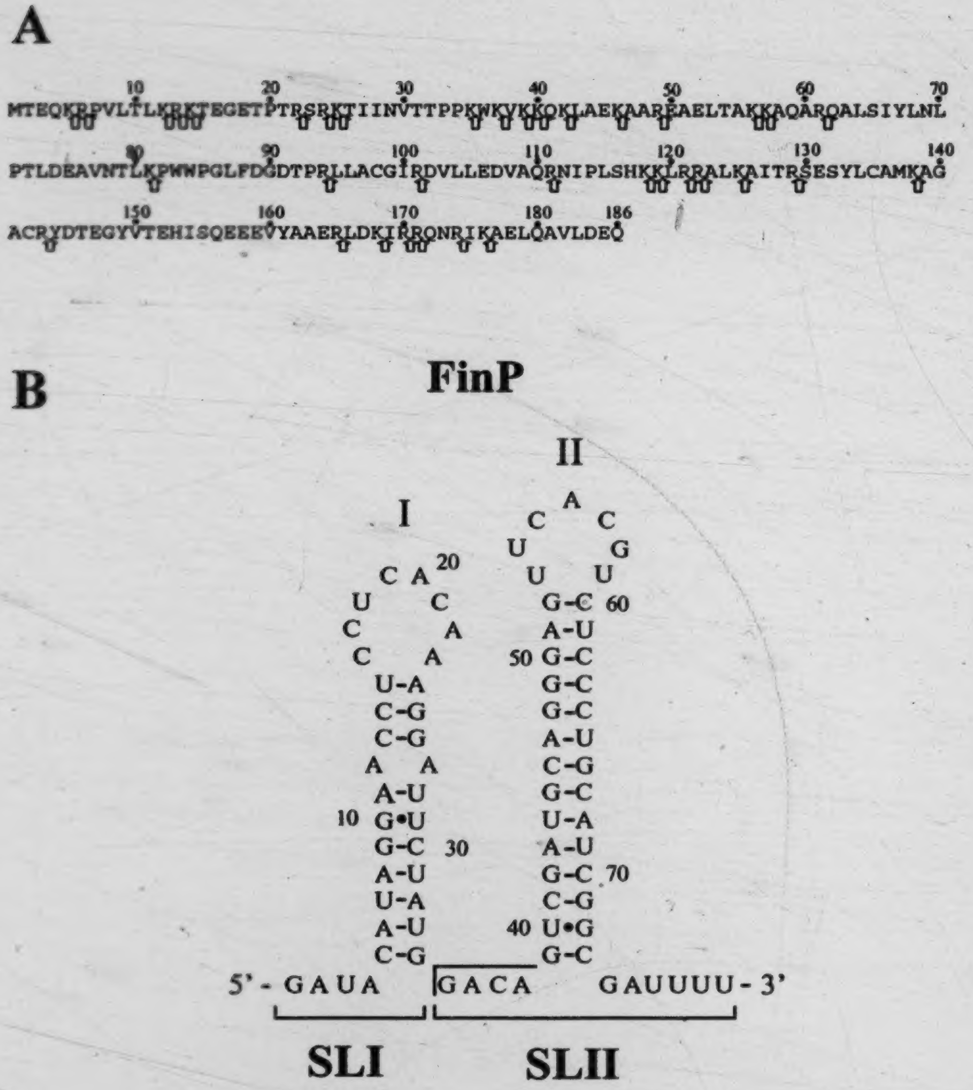

A
10
20
30
40
50
60
70
MTEQKRPVLTLKRKTEGETPTRSRKTIINVTTPPKWKVKKQKLAEKAAREAELTAKKAQARQALSIYLNL
80
90
100
110
120
130
140
PTLDEAVNTLKPWWPGLFDGDTPRLLACGIRDVLLEDVAQRNIPLSHKKLRRALKAITRSESYLCAMKAG
150
160
170
180
186
ACRYDTEGYVTEHISQEEEVYAAERLDKIRRQNRIKAELQAVLDEQ
B
FinP
I
II
20
10
30
40
50
60
70
5'-GAUA
GACA
GAUUUU-3'
SLI
SLII

Figure 3.1: Primary sequence of FinO and FinP. (A) The amino acid sequence of FinO encoded by the R6-5 plasmid is displayed using the single-letter code. Arrows indicate the positions of potential trypsin cleavage sites immediately C-terminal to lysine and arginine residues. (B) The nucleotide sequence and secondary structure of FinP are displayed. The two stem-loop regions (SLI and SLII) constituting FinP are indicated (Jerome et al., 1999).

these studies showed that FinO recognizes a specific RNA structure that is common to both FinP and *traJ* RNA. The minimal RNA target appears to be a stem-loop structure with 5' and 3' single-stranded tails. Because of this, FinO derived from one plasmid can repress the transfer of other F-like plasmids (Willetts & Maule, 1986).

The structural principles that underlie the ability of FinO to interact with RNA are currently unknown. To delimit functionally important regions of FinO, Sandercock and Frost (1998) assayed the ability of a variety of FinO deletion mutants to bind RNA, catalyze sense-anti-sense recognition, inhibit cleavage of FinP, and repress F plasmid conjugation. These studies showed that N-terminal fragments of FinO, but not a C-terminal FinO fragment, could specifically bind FinP and *traJ* RNA and catalyze their hybridization in vitro. Although these results suggested that the C-terminal region of FinO does not bind RNA, these studies demonstrated that sequences near the C-terminus are nevertheless required for stabilization of FinP and repression of F transfer in vivo. Motivated by these apparently contradictory results, we have probed the domain structure of FinO using limited proteolysis and circular dichroism spectroscopy and assayed the ability of the isolated domains to interact specifically with FinP. These experiments show that FinO is a largely helical protein that binds to individual RNA binding sites as a monomer. Surprisingly, our results demonstrate that FinO contains two separate regions that each specifically bind FinP RNA. The core of one of these regions is located at the N-terminus of the protein, between residues 26 and 61. The second region is comprised of residues 62-186, and requires sequences at the extreme C-terminus of the protein (residues 175-186) for RNA binding. Taken together with the results of Sandercock and Frost (1998), these results suggest that the C-terminal region of FinO blocks RNaseE cleavage of FinP through direct contacts with the RNA.

3.2 Results

3.2.1 RNA binding protects N- and C-terminal regions of FinO from proteolysis.

Limited proteolysis is a powerful tool to map domains of ordered secondary and tertiary structure within proteins. In general, segments of polypeptide chain that are unfolded are cleaved more readily than stably folded segments. We used trypsin to probe the structure of FinO because potential trypsin cleavage sites (lysine and arginine residues) are well dispersed throughout this highly basic protein (Figure 3.1A). To test for changes in FinO structure upon RNA binding, parallel proteolysis experiments were performed in the presence or in the absence of SLII RNA, a fragment of FinP that contains a single, high-affinity FinO binding site (Jerome & Frost, 1999). FinO protein used in these studies was overexpressed in *E. coli* and purified to near-homogeneity. SLII RNA was synthesized by *in vitro* transcription and purified by denaturing PAGE (see Experimental Procedures). Reactions were initiated by addition of trypsin to a solution of FinO or FinO in the presence of SLII RNA (Figure 3.1B) at a 4:1 molar ratio of SLII to FinO. Aliquots were taken at several time points during the reaction, and the digestion products were analyzed by SDS-PAGE and electrospray mass spectrometry (Figure 3.2). At 25 °C, FinO was digested within 360 minutes to two major trypsin-resistant fragments (Figure 3.2A). We could not separate the two species by reversed phase HPLC, and instead the mixture was analyzed by mass spectroscopy. Two major peaks were observed: one with a mass of 12 952 ± 6.5 Da, corresponding to a fragment spanning residues 62-174 (predicted mass = 12 959 Da); and the other with a mass of 12 404 ± 17 Da, corresponding to a related fragment spanning residues 62-170 (predicted molecular mass = 12 404 Da). The N-terminus of both fragments was directly determined by N-terminal amino acid sequencing of the mixture (see Experimental Procedures). In the presence of SLII, however, a larger FinO fragment was partially protected against proteolytic digestion (Figure 3.2B). A sample from a 120 minutes digestion was subjected to mass spectroscopic analysis, which revealed the molecular mass for this fragment to be 15 598 ± 8 Da, corresponding to a fragment comprising residues 50-186 (predicted molecular mass = 15 595 Da). Comparison of the earliest digestion time points in Figure 3.2A and Figure 3.2B indicates that sequences N-terminal to residue 50 are also stabilized against trypsin digestion when bound to SLII. Similar experiments, performed at 4 °C to slow the digestion, confirm this result (Figure 3.2C,D). One larger fragment, migrating slightly faster than the 22 kDa molecular mass marker, was stabilized by RNA in these experiments, although, unlike FinO(50-186) at room temperature, this fragment was also somewhat stable in the absence of RNA. Mass spectrometric analysis of a 30 minute reaction in the presence of RNA at 4 °C (Figure 3.2D) revealed that the major proteolytic species are three related fragments that comigrate in SDS-PAGE. The three fragments

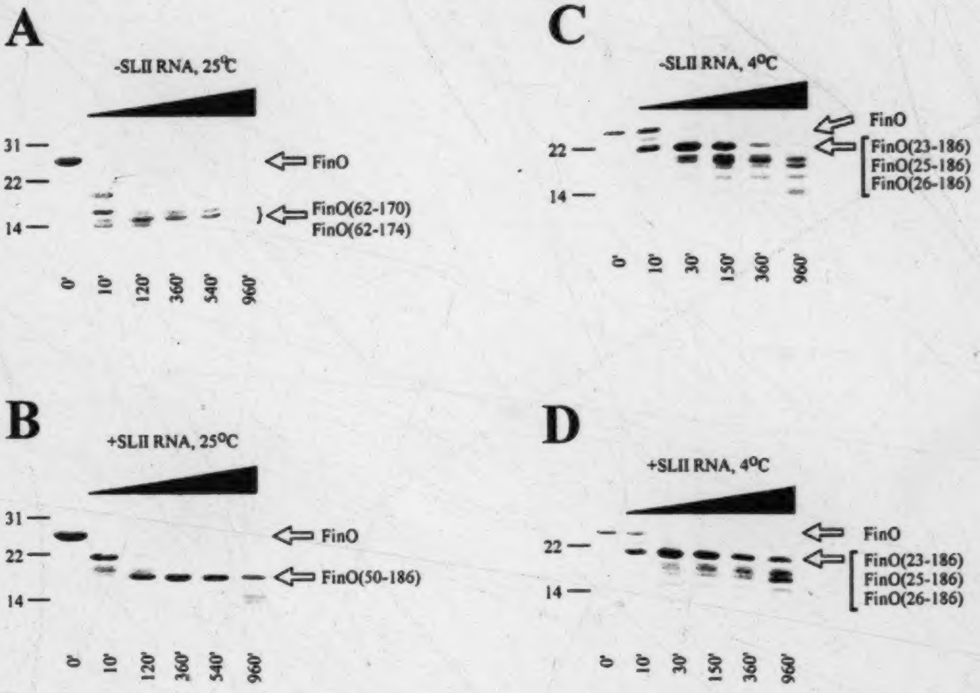


Figure 3.2: N- and C-terminal regions of FinO are protected against trypsin digestion by FinP RNA. Purified FinO was incubated either in the absence (panels **A** and **C**) or in the presence of a molar excess of SLII RNA (panels **B** and **D**) and digested with trypsin for the indicated times and temperatures. The reaction products were separated by 15% SDS-PAGE and visualized by Coomassie blue staining (see Experimental Procedures). Arrows indicate the positions of migration of FinO and several of the proteolytically derived fragments. Note that in **(C)** and **(D)** FinO(23-186), FinO(25-186), and FinO(26-186) all comigrate.

span residues 23-186 (experimentally determined molecular mass = 18 695 ± 6 Da, predicted molecular mass = 18 697 Da), 25-186 (experimentally determined molecular mass = 18 457 ± 6 Da, predicted molecular mass = 18 454 Da), and 26-186 (experimentally determined molecular mass = 18 330 ± 6 Da, predicted molecular mass = 18 326 Da). These results suggest that residues 1-61 at the N-terminus and residues 170-186 at the C-terminus of FinO are flexible, relative to the central region (residues 62-170). In complex with SLII RNA, residues 23-61 in the N-terminal region and potential trypsin cleavage sites around residues 170-174 near the C-terminus become more resistant to trypsin digestion, suggesting that these residues directly contact RNA and/or undergo a transition to a proteolytically resistant conformation in response to RNA binding.

3.2.2 Analysis of the structure of FinO by circular dichroism spectroscopy

To better understand the structures of the N-terminal, central, and C-terminal regions of FinO, we overexpressed and purified fragments of FinO corresponding to a number of the FinO trypsin fragments (Figure 3.3) and analyzed their structures by circular dichroism spectroscopy (Figure 3.4). The spectra were all measured under identical conditions to facilitate their comparison. The far-UV spectra of FinO, FinO(62-174), and FinO(62-186) are all very similar, and display double minima near 208 and 220 nm, indicating that all these proteins contain a large proportion of α-helix (Figure 3.4A). Provencher-Glöckner analysis (Provencher & Glöckner, 1981) of these spectra indicates that each fragment contains approximately 50% α-helix. In contrast, the far-UV spectrum of the N-terminal region, FinO(1-61), shows a major minimum near 200 nm and a much smaller minimum near 222 nm, indicative of random coil structure in this isolated fragment and much less helical structure. However, the difference spectrum [FinO - FinO(62-186)] (where the FinO(62-186) spectrum has been scaled to FinO) displays double minima near 208 and 222 nm, indicative of α-helical structure (Figure 3.4B). Comparison of the FinO(1-61) spectrum and the [FinO - FinO(62-186)] difference spectrum therefore suggests that the 1-61 and 62-186 domains interact in full-length FinO to induce helical structure, most likely in the N-terminal (1-61) region.

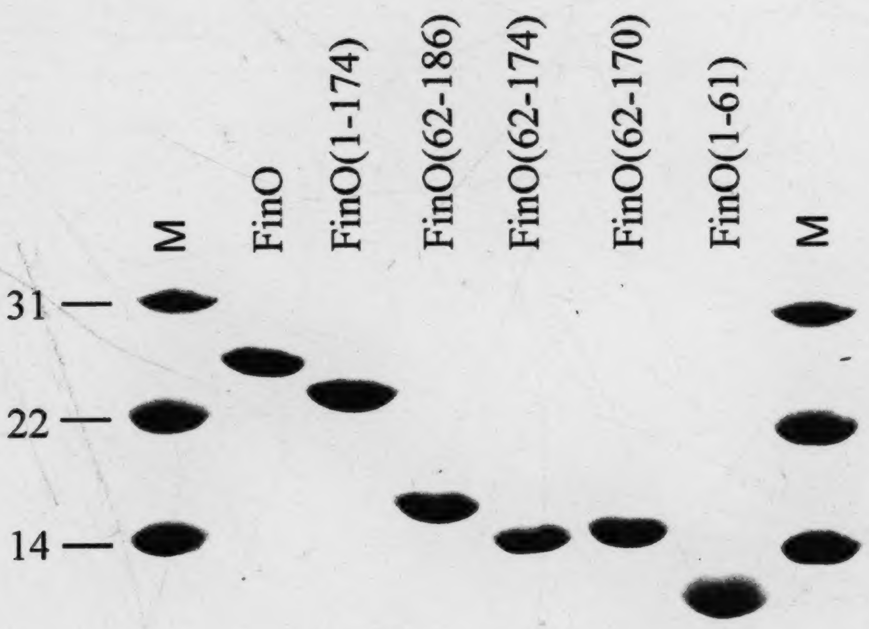


Figure 3.3: Denaturing PAGE of purified FinO and FinO fragments. FinO, FinO(62-174), FinO(62-170), FinO(1-174), FinO(62-186), and FinO(1-61) were overexpressed in *E. coli* and purified as described under Experimental Procedures. Approximately 2 μg of each purified protein was analyzed by 15% SDS-PAGE, and visualized by Coomassie blue staining.

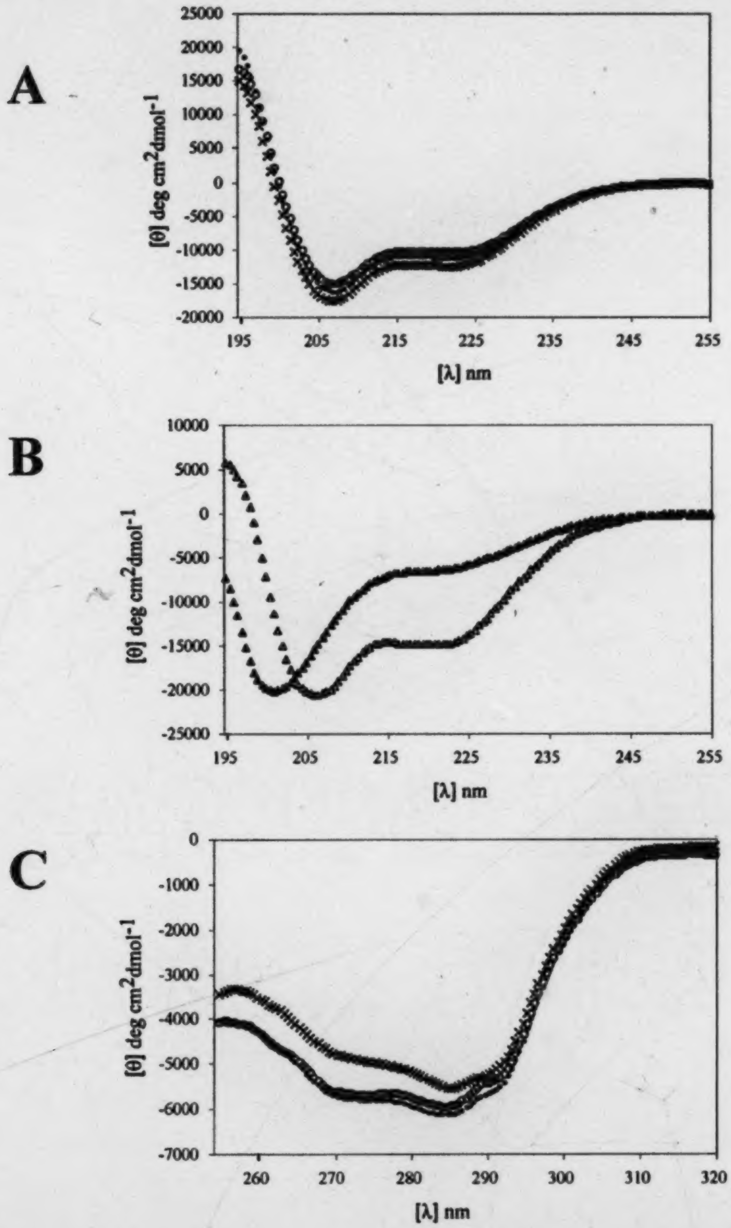


Figure 3.4: Circular dichroism spectroscopy of FinO and FinO fragments. **(A)** The far-UV CD spectra of full-length FinO (+), FinO(62-186) (●), and FinO(62-174) (O) were determined at 4 °C. **(B)** The far-UV CD spectrum of FinO(1-61) (▲) and the difference spectrum [FinO-FinO(62-186)] (Δ) were determined at 4 °C. **(C)** The near-UV CD spectra of full-length FinO (+), FinO(62-186) (●), and FinO(62-174) (O) were determined at 4 °C. Mean molar ellipticities were calculated per number of aromatic residues to allow for comparison.

Comparison of the far-UV CD spectra of FinO, FinO(62-186), and FinO(62-174) indicates that deletion of either the N- or the C-terminal region does not greatly affect the secondary structure of the central domain. To determine if the N- and C-terminal regions have any effect on the tertiary structure of the central domain, we also measured near-UV CD spectra of these proteins (Figure 3.4C). Circular dichroism in the near-UV region (255-320 nm) is largely affected by the packing interactions of aromatic residues. Aromatic residues that are packed within a hydrophobic protein core tend to yield complex near-UV spectra with either positive or negative ellipticity. In contrast, protein structures that lack tertiary packing interactions give very little signal in near-UV CD spectra (Woody, 1995). FinO, FinO(62-186), and FinO(62-174) have seven aromatic residues in common, with an additional N-terminal tryptophan in full-length FinO. The similarity in the number of aromatic residues in these fragments allows a comparative analysis of the spectra of these proteins (Figure 3.4C). The overall pattern and the magnitude of the peaks are very similar for all three proteins, indicating that the packing of the common aromatic residues is similar in all three proteins. Taken together with the far-UV CD results, we can conclude that neither residues 1-61 nor residues 175-186 exert a significant effect on the secondary or tertiary structure of the central domain of FinO (residues 62-174).

3.2.3 The N- and C-terminal regions of FinO each specifically bind FinP

Our proteolytic studies suggested that the N-terminal region (residues 1-61) and the C-terminal region (residues 175-186) contact FinP. To test this hypothesis more directly, we measured the ability of purified FinO and its fragments (Figure 3.3) to bind FinP RNA in an electrophoretic mobility shift assay (EMSA) (see Experimental Procedures). To determine if the observed interactions were specific for FinP, we also performed parallel binding reactions in the presence of a 1000-fold molar excess of total *E. coli* tRNA as nonspecific competitor. The results of several representative experiments are shown in Figure 3.5. Each protein fragment that contains either the N-terminal region [FinO (Figure 3.5A), FinO(1-61) (Figure 3.5B), or FinO(1-174) (Figure 3.5C)] or the C-terminal region [FinO (Figure 3.5A) or FinO(62-186) (Figure 3.5D)] is able to bind FinP even in the presence of competitor tRNA, giving rise to several species of reduced

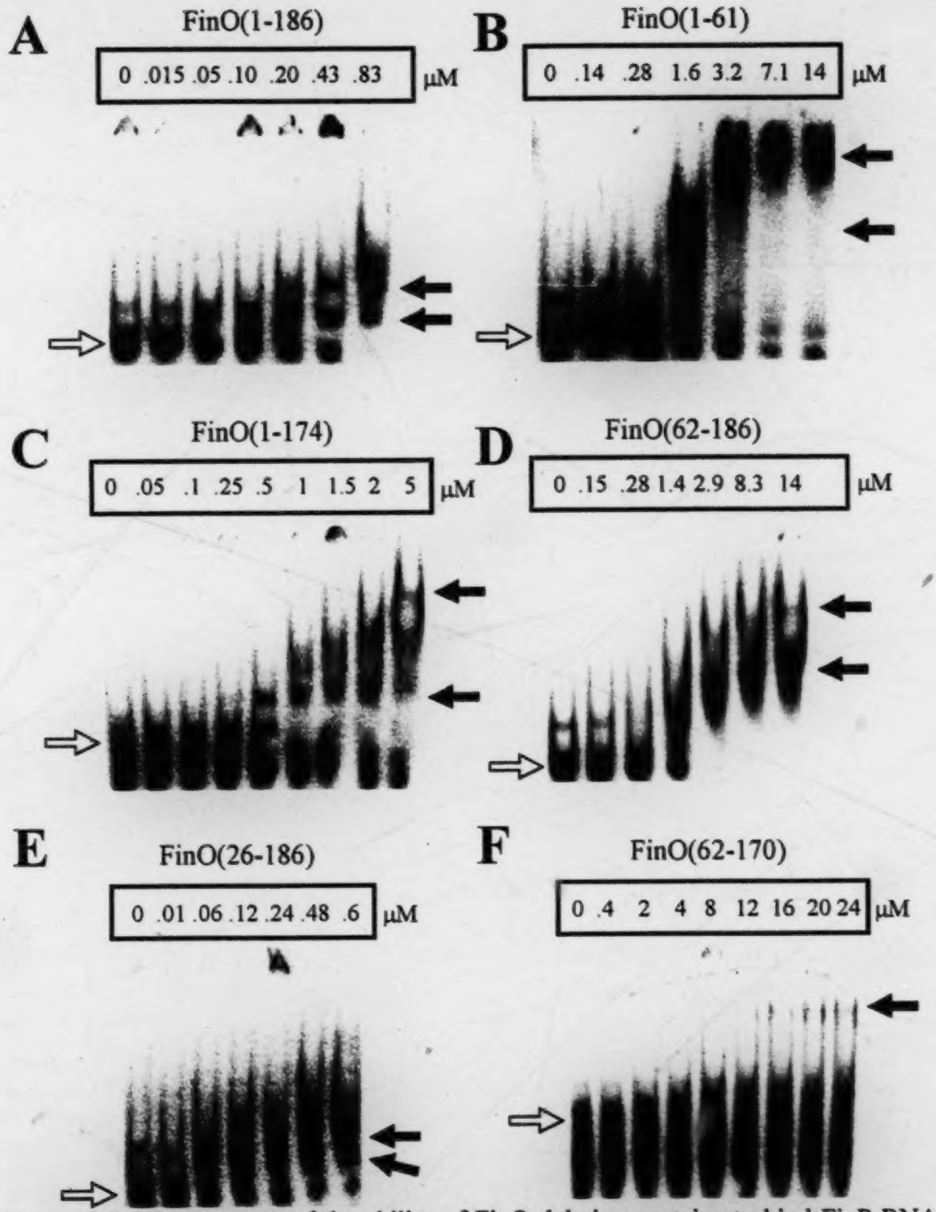

A
FinO(1-186)
0 .015 .05 .10 .20 .43 .83 μM
B
FinO(1-61)
0 .14 .28 1.6 3.2 7.1 14 μM
C
FinO(1-174)
0 .05 .1 .25 .5 1 1.5 2 5 μM
D
FinO(62-186)
0 .15 .28 1.4 2.9 8.3 14 μM
E
FinO(26-186)
0 .01 .06 .12 .24 .48 .6 μM
F
FinO(62-170)
0 .4 2 4 8 12 16 20 24 μM

Figure 3.5: Measurement of the ability of FinO deletion proteins to bind FinP RNA by EMSA. In each case, FinO protein or the indicated fragment was incubated with 7.5 fmol of 32 P-labeled FinP RNA. Bound (closed arrows) and unbound (open arrows) FinP shown in each gel were resolved by electrophoresis on 5% polyacrylamide gels and visualized by phosphorimaging. The concentration of protein present in each binding reaction is shown above each lane. In panels (**A**)-(**E**), all binding reactions were performed in the presence of a 1000-fold molar excess of total *E. coli* tRNA as nonspecific competitor; in panel (**F**), no competitor was used. These experiments were performed by Michael Gubbins.

mobility relative to free FinP. The presence of multiple retarded species suggests that there are several FinO binding sites in FinP (van Biesen & Frost, 1994; Jerome & Frost, 1999). In contrast, fragments lacking both the N- and C-terminal regions [FinO(62-174) (data not shown) and FinO(62-170) (Figure 3.5F)] bind only weakly to FinP.

In order to quantitate the relative contributions of the N- and C-terminal domains to FinP binding, we calculated the apparent RNA binding constant ($K_{a,app}$) for each fragment in the presence or the absence of competitor tRNA (Table 3.1) (see Experimental Procedures). In the absence of tRNA competitor, full-length FinO was found to bind FinP with an apparent binding constant of 2.0×10^7 M^{-1}, similar to the binding constant determined previously for FinO fused to GST (Sandercock & Frost, 1998; Jerome & Frost, 1999). Deletion of residues 1-61 led to a ~5-fold reduction in the binding affinity while deletion of residues 175-186 had a somewhat greater effect, reducing the FinP binding affinity ~10-fold. FinO(62-174) and FinO(62-170), in which both the N- and C-terminal domains have been deleted, bind FinP too weakly to allow accurate quantitation of the their binding affinities. However, the affinity of these fragments for FinP appears to be reduced at least 1000-fold compared to full-length FinO. Thus, the central domain, in isolation, plays only a very minor role in RNA binding. While the binding experiments performed in the presence of competitor indicate that the N- and C terminal domains both bind FinP specifically, a quantitative analysis of the degree of competition suggests that the specificity of the individual domains may be slightly different.

While the apparent affinity of full-length FinO for FinP decreases about 3.9-fold when challenged with tRNA competitor, the apparent affinities of FinO(1-61) and FinO(1-174) for FinP are less affected, decreasing 1.5- and 2.5-fold, respectively, when challenged with the same molar excess of competitor. In contrast, the binding affinity of the C-terminal fragment, FinO(62-186), is more affected by tRNA challenge, decreasing 5.2-fold. Thus, the N-terminal domain may bind FinP somewhat more specifically than the C-terminal domain.

The proteolysis experiments suggested that residues 23-61 were protected against trypsin digestion when bound by RNA but protection of sites closer to the N-terminus could not be detected. To test the role of these more N-terminal residues in RNA binding, we also

3.1: Apparent K_a of the FinO deletion proteins binding to FinP[1]

FinO fragment	K_a (M^{-1}) No competitor[3]	+ competitor[3]
whole FinO	$(2.0 \pm 0.5) \times 10^7$	$(5.1 \pm 0.3) \times 10^6$
1-174	$(2.7 \pm 1.1) \times 10^6$	$(1.1 \pm 0.2) \times 10^6$
26-186	$(3.2 \pm 0.3) \times 10^7$	$(7.3 \pm 1.8) \times 10^6$
62-186	$(4.2 \pm 1.0) \times 10^6$	$(8.1 \pm 1.0) \times 10^5$
1-61	$(2.1 \pm 0.7) \times 10^6$	$(1.4 \pm 0.1) \times 10^6$
62-170[2]	1×10^4	ND[4]
62-174[2]	1×10^4	ND

---

[1] These experiments were performed by Mike Gubbins

[2] Ka was estimated due to the inability of FinO deletion protein to bind more than approximately 35-40% of the ^{32}P-labeled FinP present in a given EMSA except at extremely high protein concentrations

[3] Competitor was E. coli total tRNA in a 1000-fold molar excess compared to FinP

[4] ND: not determined

determined the affinity of purified FinO(26-186) for FinP (Figure 3.5E). The binding constant of this protein is essentially identical to that of FinO, both in the presence and in the absence of tRNA competitor. Therefore, we conclude the N-terminal RNA binding domain lies within residues 26-61.

3.2.4 FinO binds RNA as a monomer

FinO might facilitate *traJ*-FinP interactions by simply bringing the two RNAs together, initially through protein-protein interactions between two FinO molecules, one bound to TraJ and the other to FinP. To test this hypothesis, we first employed equilibrium analytical ultracentrifugation to probe the oligomeric state of free FinO in solution (data not shown). These experiments indicated that FinO exists primarily as a monomer at concentrations similar to *in vivo* levels (~1-10 μM). However, these experiments also demonstrated that FinO aggregates at much higher concentrations (~1 mM), leaving open the possibility that FinO might oligomerize when bound to RNA.

We next used EMSA to determine the oligomeric state of FinO when bound to the single binding site present in SLII RNA (Figure 3.6). A similar method was first used to show that the bZIP DNA binding protein GCN4 interacts with DNA as a dimer (Hope & Struhl, 1987). In this experiment, we bound FinO, GST-FinO, or a mixture of these two proteins to SLII and separated the resulting complexes by nondenaturing PAGE (Figure 3.6). As previously shown (Jerome & Frost, 1999), GST-FinO binds this RNA to give a single protein-RNA complex with reduced mobility compared to free SLII. Likewise, FinO also binds SLII to give a single protein RNA-complex. Due to the difference in molecular mass between FinO and GST-FinO, the FinO-SLII complex migrates faster than the GST-FinO- SLII complex. To determine the stoichiometry of binding, we premixed GST-FinO and FinO, added SLII, and separated the resulting mixture by nondenaturing PAGE. We reasoned that if FinO binds SLII as a dimer or higher order aggregate, we would observe not only GST-FinO-SLII and FinO-SLII complexes, but also heteromeric GST-FinO-FinO complexes bound to SLII. The heteromeric complexes would be expected to migrate in the EMSA between the positions of the GST-FinO-SLII and FinO-SLII complexes. However, as shown in Figure 3.6, only GST-FinO-SLII and FinO- SLII complexes are observed. The fact that FinO and GST-FinO both bind FinP

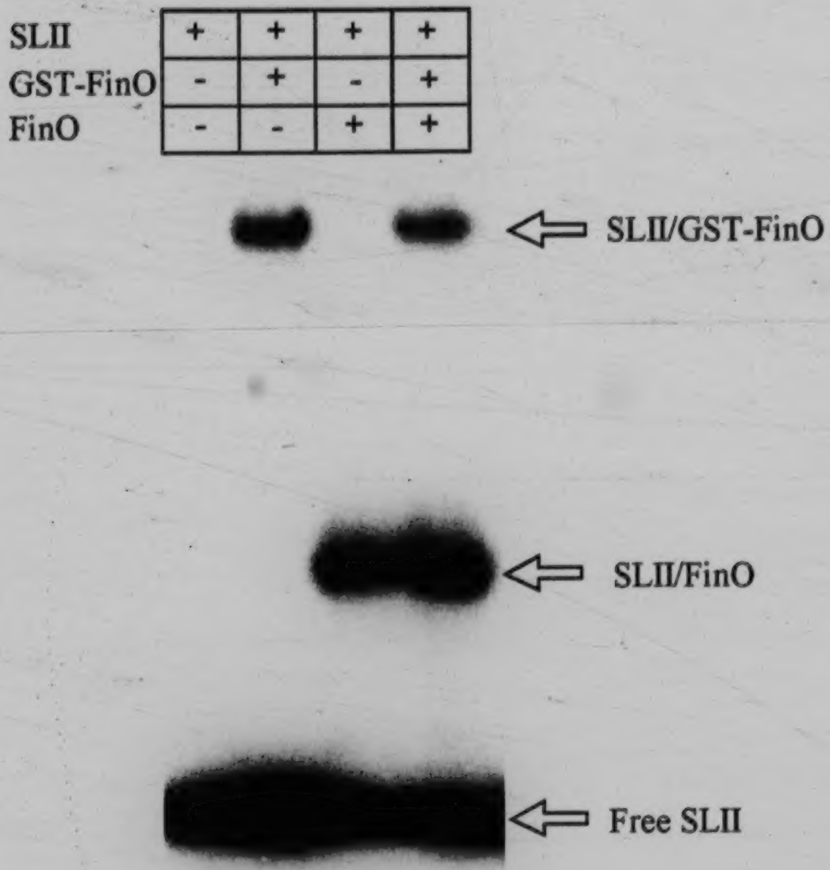


Figure 3.6: EMSA demonstrates that FinO binds FinP as a monomer. ^{32}P-labeled SLII RNA was incubated with FinO and/or GST-FinO, as indicated at the top of the figure, and the resulting protein-RNA complexes were analyzed by nondenaturing gel electrophoresis as described under Experimental Procedures. Arrows indicate the positions of the SLII/FinO and SLII/GST-FinO complexes, as well as free SLII.

with almost identical affinities [~(2-5) x 10^7 M^{-1}, Table 3.1 and Sandercock & Frost, 1998] suggests that fusion to GST does not significantly affect the oligomeric state of FinO when bound to RNA. We therefore conclude that, under these conditions, FinO binds the SLII region of FinP as a monomer.

3.3 Discussion

Our results indicate that FinO contains two distinct RNA binding regions, which can bind RNA independently of one another. One region is located within the N-terminal third of the protein, between residues 26 and 61, whereas the second region extends from residues 62 to 186. We have not delimited the precise boundaries of the C-terminal RNA binding region, however, two lines of evidence suggest that its primary RNA contact surface is located in the vicinity of a cluster of positively charged residues (165 to 176) near the C-terminus of the protein (Figure 3.1A). First, FinP binding protects FinO from trypsin cleavage at residues 170 and 174, suggesting that FinP binding sterically blocks access of the protease to this region of FinO. Second, deletion of residues 175-186 from FinO(62-186) almost completely abolishes the ability of this protein to interact with FinP. The similarity of the CD spectra of FinO(62-174) and FinO(62-186) suggests that this is not an indirect effect caused by a conformational change in residues 62-174 upon deletion of residues 175-186 (Figure 3.4). We have not mapped the precise N-terminal boundary of the C-terminal domain; however, our proteolytic mapping studies suggest that this boundary lies somewhere between residue 62 and the next potential trypsin cleavage site at residue 81 (Figure 3.1A). Taken together with our results, the previous finding that a 37 amino acid C-terminal fragment of FinO does not bind FinP (Sandercock & Frost, 1998) suggests that, while this region may contain an essential RNA contact surface, it requires the entire 62-186 domain to interact with FinP.

Sandercock and Frost (1998) previously showed that residues 161 to 186 at the C-terminus of FinO are required to protect FinP from digestion by cellular RNases. RNaseE appears to be the enzyme responsible for FinP degradation *in vivo*, cleaving FinP at a specific site located within a single-stranded region that links the two stem-loop structures (Jerome & Frost, 1999). This single-stranded region is also required for high-affinity binding of FinP by FinO (Jerome & Frost, 1999). It therefore is likely that

FinO(62-186) contacts this single-stranded region, and may sterically block cleavage of this strand by RNaseE.

Full-length FinO or FinO(26-186), both of which contain the N- and C-terminal RNA binding regions, yield complexes with FinP that have distinct, defined electrophoretic mobilities, giving rise to relatively sharp bands in EMSA (Figure 3.5). In contrast, the fragments which contain only one of the RNA binding regions form complexes with FinP that give much more diffuse bands, and retard the RNA to a greater degree than either FinO or FinO(26-186). This behavior may be explained by nonspecific aggregation of FinO(1-61), FinO(1-174), and FinO(62-186) on the RNA. The diffuse nature of the bands observed with these fragments might also indicate that a significant proportion of these complexes dissociate during electrophoresis. These observations suggest that both RNA binding domains are required to form a kinetically stable complex with FinP, compared to fragments containing only one of the RNA binding regions.

Perhaps the most intriguing aspect of FinO function is the ability of the protein to facilitate sense-antisense interactions between FinP and *traJ* mRNA. The ROM protein of ColE1 performs an analogous function, facilitating the recognition of two complementary RNAs, RNAI and RNAII (Predki *et al.*, 1995). In both the F and ColE1 systems, initial RNA recognition most likely occurs through base-pairing interactions between complementary loops to form "kissing" complexes. While ROM directly interacts with the kissing complex and facilitates its conversion to duplex, FinO instead binds to the individual RNAs (Jerome & Frost, 1999). How, then, does FinO facilitate FinP-*traJ* mRNA recognition? One mechanism could involve FinO bringing together FinP and *traJ* via protein-protein interactions between FinO molecules bound to separate RNAs, which have formed the kissing complex. Our results, however, show that FinO binds an isolated, single RNA target as a monomer. Therefore, if FinO does facilitate FinP-*traJ* mRNA recognition through protein-protein interactions, these interactions must be extremely weak in the absence of both RNA targets. Alternatively, a FinO monomer may stabilize the kissing complex, or FinO-dependent RNA annealing may be facilitated by some other mechanism. For example, FinO might destabilize intramolecular base-pairing within FinP and *traJ* mRNA, allowing complementary base pairs to form between these two RNAs. In chapter 6 we present evidence that FinO does possess unwinding activity that is required for promotion of duplex formation.

FinO bears no sequence similarity to other known RNA binding proteins. While our studies clearly show that there are two separate regions of FinO that contact RNA, we still do not know if these regions form two independent structural domains, or if the two regions come together in the intact protein to form a single contiguous RNA contact surface. High-resolution structural studies of FinO, alone and bound to minimal RNA targets, will be required to answer this question and reveal, at a fundamental level, how this protein recognizes RNA in a structure-specific manner.

Chapter 4

Crystal structure of the bacterial conjugation repressor FinO*

* A version of this chapter is published in; Nature Structural Biology, Vol. 7, No. 7, 2000, Alexandru F. Ghetu, Michael J. Gubbins, Laura S. Frost and J. N. Mark Glover.

4.1 Introduction

The 35 kb transfer (*tra*) operon of the F-plasmid encodes the proteins responsible for conjugative transfer of this plasmid from host to recipient *Escherichia coli* cells (Frost *et al.*, 1994). Expression of the *tra* operon is repressed by the two component FinOP system that includes the 186 residue, basic protein FinO, and the ~79 nucleotide RNA, FinP (Finnegan & Willetts, 1972). FinP contains two stem-loop structures (SLI and SLII) and is complementary to the untranslated leader of *traJ* that encodes the primary transcriptional activator of the *tra* operon (Mullineaux & Willetts, 1985). FinP is thought to interact with *traJ* mRNA to occlude its ribosome binding site, blocking *traJ* translation and thereby inhibiting transcription of the *tra* operon (van Biesen & Frost, 1994).

Unprotected FinP is rapidly degraded by the cellular endonuclease RNase E and is, therefore, ineffective in repressing *traJ* translation (Jerome *et al.*, 1999). FinO, the second component of the inhibition system, protects FinP against degradation. It binds to FinP and sterically blocks access to the RNase E cleavage site (Jerome *et al.*, 1999). FinO also binds to the complementary stem-loop structures in *traJ* mRNA and promotes duplex formation between FinP and *traJ* RNA *in vitro* (Lee *et al.*, 1992; van Biesen *et al.*, 1993; Koraimann *et al.*, 1996).

We showed that amino acids 26–61 and the remaining C-terminal residues (62–186) of FinO constitute two independent RNA binding regions (Ghetu *et al.*, 1999, also see chapter 3). The C-terminal region is also important for conferring RNase E resistance to FinP (Sandercock & Frost, 1998). Within this region, residues 62–174 comprise a proteolytically stable domain that requires the rest of the C-terminal region (residues 175–186), or the N-terminal RNA binding region, for it to interact with RNA (Ghetu *et al.*, 1999, also see chapter 3). Here we present the crystal structure of a fragment of FinO (residues 26–186) that contains both RNA recognition regions. Using this structure, we have built a testable model of a complex of FinO bound to RNA comprising a minimal binding sequence and suggest a means by which FinO could facilitate FinP–*traJ* RNA interactions.

4.2 Results

4.2.1 Structure determination

FinO protein was overexpressed and purified from *E. coli* as described in Experimental Procedures. Attempts to crystallize full length FinO from the F-like plasmid R6-5 were unsuccessful. We have shown that FinO lacking only the N-terminal 25 amino acids (FinO(26–186)) comprises a proteolytically stable fragment at 4 °C that binds FinP RNA with the same affinity as the full length protein (Chapter 3). FinO(26–186) crystallized at 4 °C in the space group $P2_12_12_1$ ($a = 37.57$ Å, $b = 38.73$ Å, $c = 145.42$ Å) with one FinO molecule per asymmetric unit. The crystals diffracted X-rays weakly to 3.0 Å on a rotating anode X-ray source.

The structure of FinO(26–186) was solved by multiple wave-length anomalous dispersion (MAD) methods using selenomethionine substituted protein. Native FinO has only one methionine site at which selenomethionine can be introduced. To increase the phasing power of the data collected from selenomethionine substituted crystals, we mutated leucines 96 and 124 to methionines (see Experimental Procedures). Crystals of the selenomethionine substituted double mutant FinO(26–186)L96,124M had the same unit cell dimensions as the wild type. Electrophoretic gel mobility shift analysis demonstrated that FinO(26–186)L96,124M binds FinP RNA with the same affinity as wild type FinO and FinO(26–186) (data not shown). A three-wavelength MAD data set was collected at beamline X12C at the National Synchrotron Light Source (NSLS). Solvent flattened, MAD phased electron density maps calculated to 2.6 Å resolution revealed excellent main and side chain density for residues 54–184 (Figure 4.1A). Residues 33 to 53 adopted a solvent exposed helical conformation with generally poorer electron density than the rest of the molecule. Using this map, a model of FinO was built and partially refined. We subsequently collected diffraction data to 2.0 Å resolution from a crystal of the single point mutant, FinO(26–186)L124M. This data set was used in the final stages of refinement. Comparison of electron density maps calculated with the single and double mutant data indicate that the Leu to Met mutations have essentially no effect on the structure of FinO(26–186).

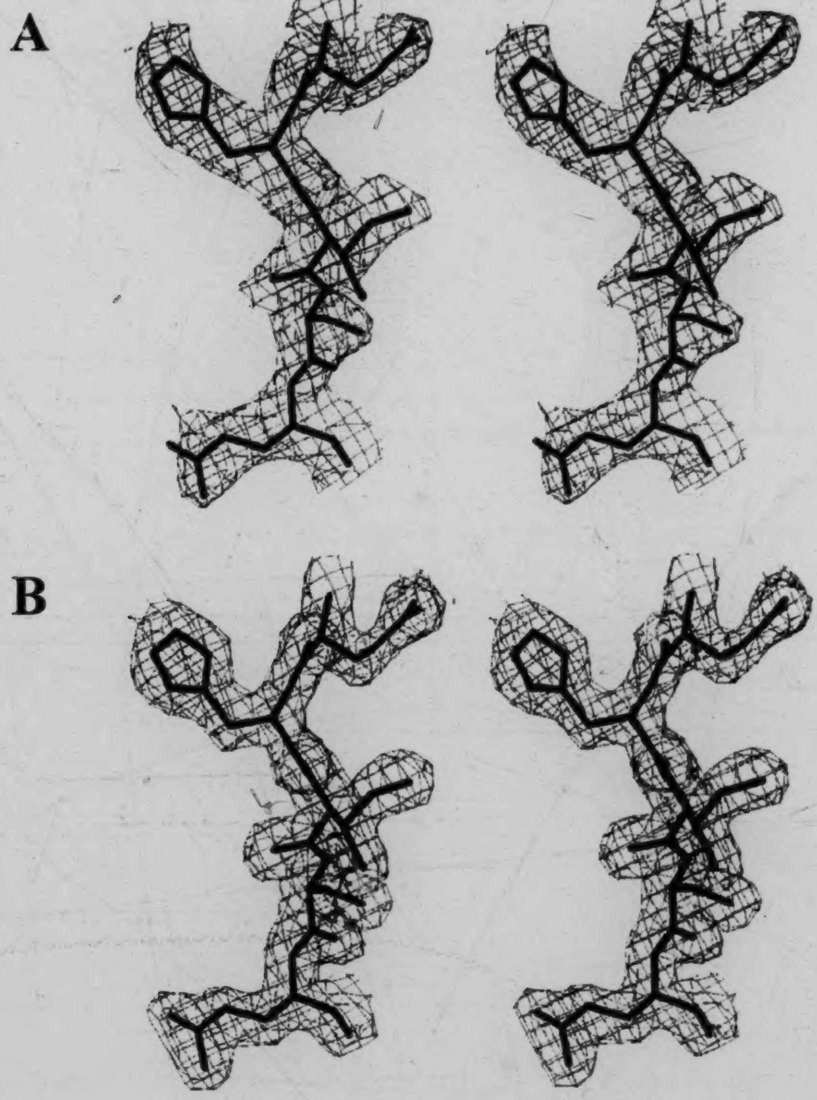


Figure 4.1: Stereo view of electron density maps around residues 152–156 with the final 2.0 Å refined model superimposed. **(A)** Density modified, 2.6 Å resolution MAD phased experimental map contoured at 1.0σ. **(B)** 2|Fo| - |Fc| map at 2.0 Å resolution using phases calculated from the final, refined model and contoured at 1.6σ.

Our current model of FinO(26–186)L124M contains residues 33–184 (Figure 4.2), and 212 water molecules. The model has been refined to an R-factor of 19.7% (R_{free} of 22.4%) using all data (20–2.0 Å) and has good stereochemistry (Table 4.1). A section of the 2|Fo| – |Fc| electron density map, phased with the final refined model, is shown in Figure 4.1B.

4.2.2 Structure description

Overall, FinO(26–186) is a largely helical, elongated molecule with a maximum length of ~80 Å (Figures 4.3, 4.4, 4.5A and 4.6). Searches of the protein structure database using either the DALI (Holm & Sander, 1993) or VAST (Madej *et al.*, 1995) web servers suggest that FinO adopts a novel protein fold. The structure is reminiscent of a right-handed fist, with an extended index finger and thumb that touches the index finger near its base. The index finger corresponds to a 45 Å long α–helix (α1) consisting of residues 33–65, which ends with a single turn of 3_{10}-helix at its C-terminus (residues 66–68). The thumb corresponds to a C-terminal helix (α6) consisting of residues 155–184. These two terminal helices interact with each other along a stretch of three helical turns. An additional 30 Å of the α1 helix is completely solvent exposed and extends past the C-terminus of FinO. The central region of FinO consists of five α-helices, including part of the C-terminal helix and two β-hairpins, β1-β2 and β3-β4. A distinct feature of FinO is the presence of a hole, ~5 Å in diameter, formed between the central region and the N-terminal and C-terminal helices.

4.2.3 FinO conformational flexibility

Our proteolytic mapping study (Chapter 3) suggested that helix 1, and, to a lesser extent, the C-terminal 17 residues in helix 6, are more flexible than the central region (residues 62–174). An analysis of the main chain temperature factors (B-factors) throughout FinO confirms this finding (Figure 4.5B). The main chain B-factors are lowest within the central region of FinO but rise dramatically in helix 1 as it emerges from the central region near residue 60, and near the C-terminus of helix 6 after residue 174.

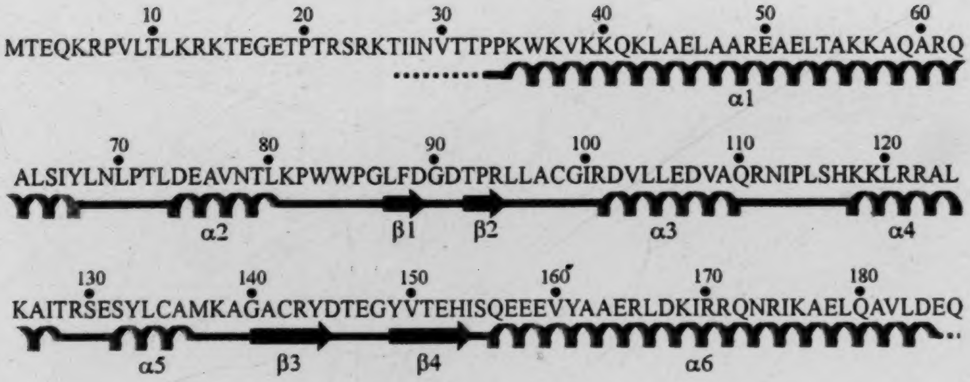

10
20
30
40
50
60
MTEQKRPVLTLKRKTEGETPTRSRKTIINVTTPPKWKVKKQKLAELAAREAELTAKKAQARQ
α1
70
80
90
100
110
120
ALSIYLNLPTLDEAVNTLKPWWPGLFDGDTPRLLACGIRDVLLEDVAQRNIPLSHKKLRRAL
α2
β1
β2
α3
α4
130
140
150
160
170
180
KAITRSESYLCAMKAGACRYDTEGYVTEHISQEEEVYAAERLDKIRRQNRIKAELQAVLDEQ
α5
β3
β4
α6

Figure 4.2: The primary sequence and secondary structure elements of FinO. Amino acids are represented in the single letter amino acid code. Secondary structure motifs are labeled below the sequence and are colored as follows: α-helix, cyan; β-strands, blue; regions of random coil, red; 3_{10}-helix, orange.

Table 4.1: X-ray data collection, phasing and refinement statistics

Data set	High resolution FinO(26-186)L124M	MAD data collection Se-FinO(26-186)L96,124M		
		λ1	λ2	λ3
Wavelength (Å)	0.9574	0.9790	0.9786	0.9640
Resolution (final shell, Å)	2.0	2.6	2.6	2.6
Observations	109,290	58,397	58,355	58,307
Unique reflections	14,726	6,841	6,839	6,832
Data coverage (%)[1]	97.3	97.2	97.2	97.2
R_{merge} total / final shell (%)[2]	4.7/11.0	6.1/10.7	6.6/11.4	6.2/11.4
Phasing (20.0-2.0 Å)				
Overall figure of merit	0.65 (SOLVE; Terwilliger, 1987)			

Refinement Statistics (20-2.0Å)			
R_{cryst}[3] total / final shell (%)	19.7 / 22.4	Average B-factor (Å^2)	24.1
R_{free}[4] total / final shell (%)	22.4 / 23.4	Number of atoms	
R.m.s. deviations		Protein	1,219
Bond lengths (Å)	0.009	Water	212
Bond angle (°)	1.35	Ramachandran analysis[5]	
		Most favored	96.4%
		Additionally allowed	3.6%

---

[1] The highest resolution shell for FinO(26-186)L124M was 2.05-2.00 Å and for Se-FinO (26-186)L96,124M it was 2.69-2.60 Å.

[2] $R_{merge} = \sum |\langle I\rangle - I| / \sum \langle I\rangle$.

[3] $R_{cryst} = \sum ||Fo| - |Fc|| / \sum |Fo|$.

[4] R_{free} calculated with 10% of all reflections excluded from refinement stages using high resolution data.

[5] Ramachandran analysis was performed with the program Procheck (Laskowsky *et al*., 1993)

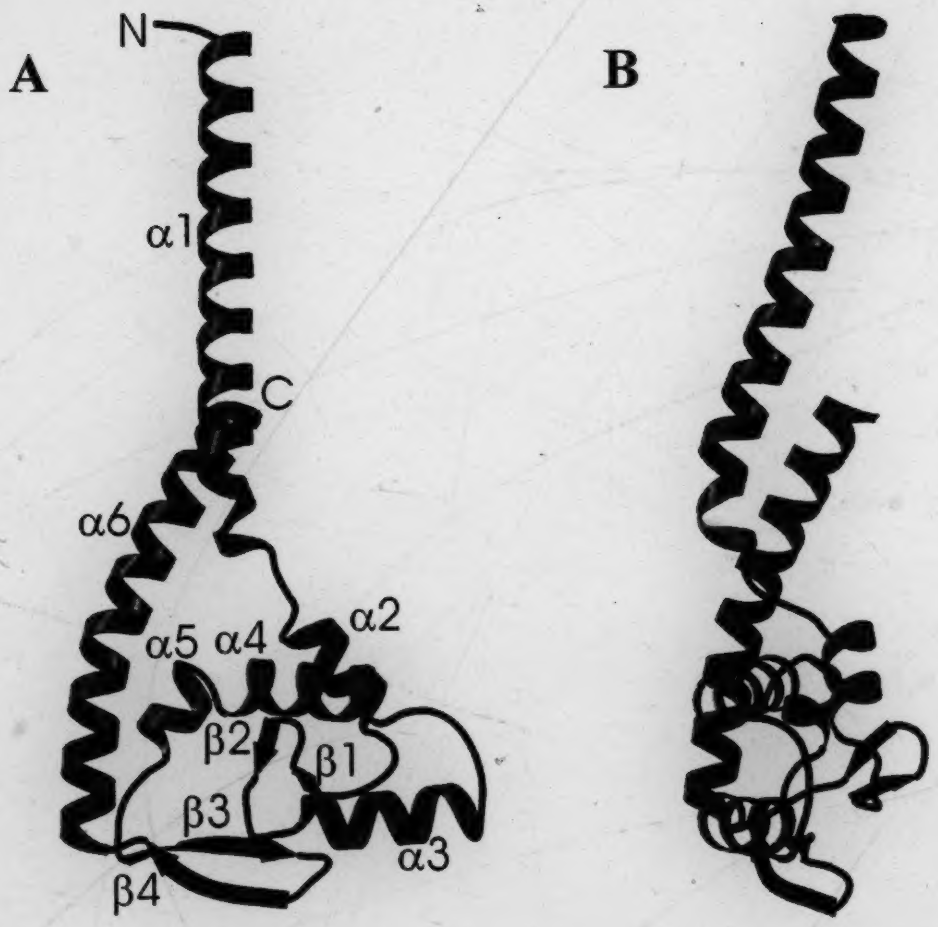


Figure 4.3: Ribbon diagrams of FinO. (**A**), The secondary structure motifs are numbered as in figure 4.2 and the N-terminus and C-terminus are labeled. (**B**), FinO rotated 90° about a vertical axis with respect to (**A**). Secondary structure motifs in (**A**) and (**B**) are colored as in figure 4.2.

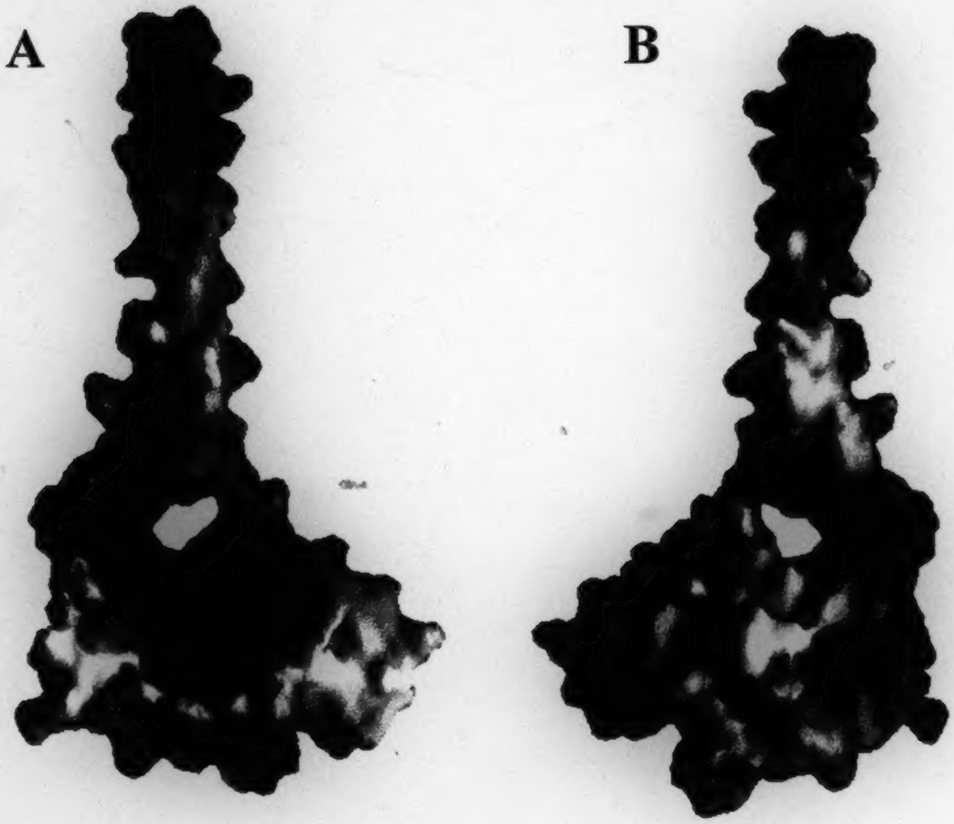


Figure 4.4: Molecular surface representation of FinO, colored by electrostatic potential. Increasing saturation of positive and negative potential is represented by blue and red, respectively. **(A)**, Protein is in the same orientation as in panel (A) of Figure 4-3. **(B)**, Electrostatic potential surface of FinO, with the protein rotated 180° with respect to **(A)**.

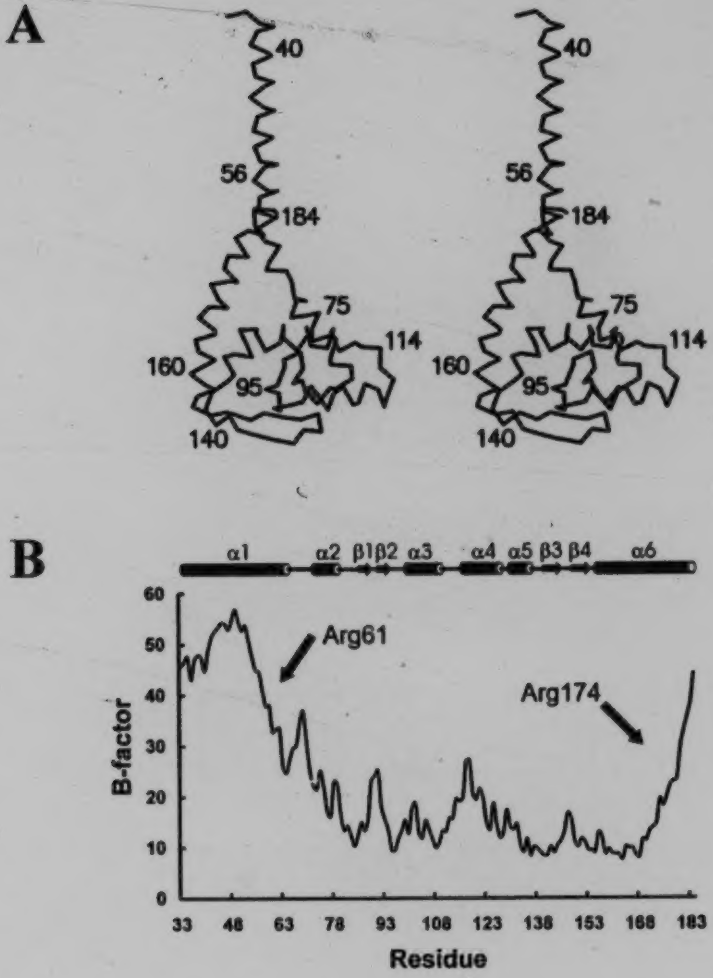


Figure 4-5: Trace of FinO main chain Cα atoms and their B-factors. **(A)**, FinO trace is shown in stereo view and the position of approximately every 20th amino acid is indicated on the diagram. **(B)**, Plot of the final refined B-factor values for the main-chain C atoms of FinO, with the secondary structure of FinO is indicated above the graph.

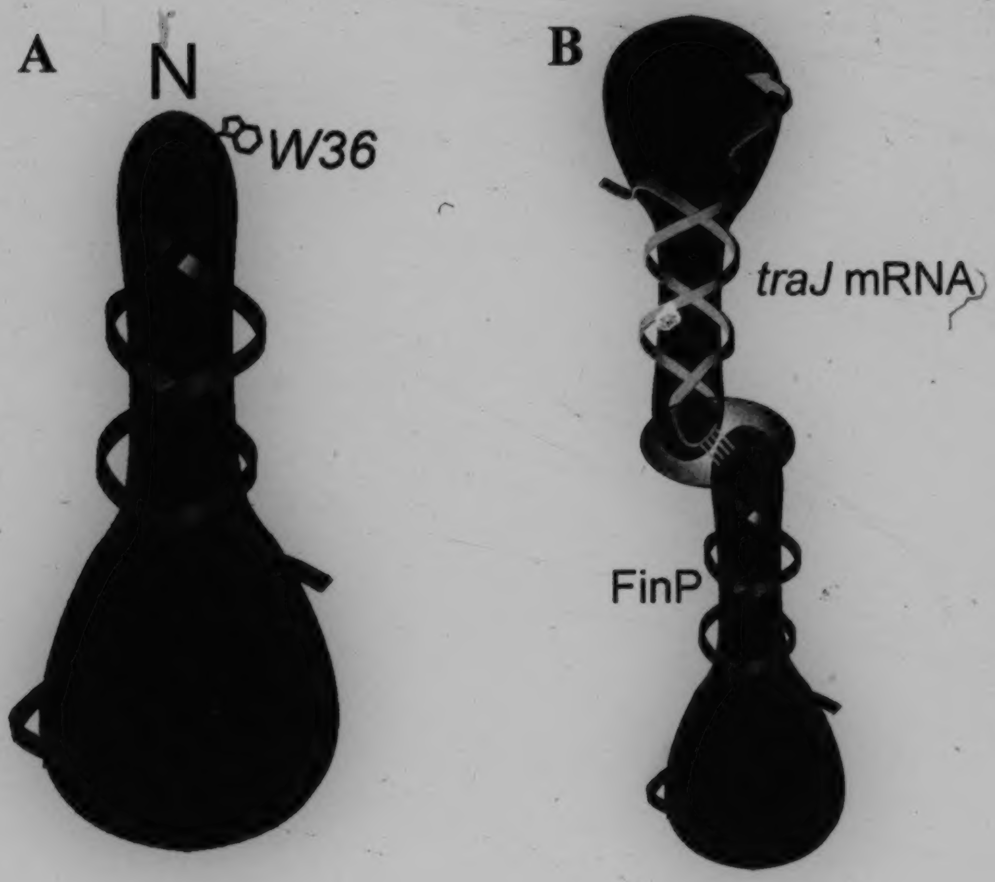


Figure 4-6: FinO–RNA interactions. (**A**), Hypothetical model of FinO bound to SLII of FinP. The Trp 36 (W36) side chain is displayed in purple. (**B**), Possible model for a FinO stabilized kissing complex formed between FinP SLII and the complementary stem-loop structure of *traJ* mRNA. The N-terminal regions of FinO are predicted to interact with and stabilize RNA loop–loop interactions in the region highlighted in green. FinO is colored blue while the RNA backbones of SLII and *traJ* mRNA are displayed as yellow and orange ribbons, respectively.

4.2.4 FinO-RNA interactions

Studies have revealed that FinO contains two, separable, RNA binding regions, one encompassing residues 26–61 and the other consisting of residues 62–186 (Chapter 3). Each region binds FinP specifically, albeit with 5- to 10-fold lower affinity than full length FinO or FinO(26–186). An electrostatic potential surface calculated for FinO(26–186) reveals positively charged regions that are likely to contact RNA directly (Figure 4.4A,B). Residues near the N-terminus of helix 1 comprise one positively charged surface in the structure. A second large positively charged surface extends across one face of the central domain, and is composed of Lys and Arg residues from helices 4 and 6. Proteolytic mapping and direct RNA binding measurements of FinO deletion mutants have shown that these regions in helices 1 and 6 do contact RNA, although a direct role for helix 4 in RNA binding has yet to be established (Chapter 3).

A quantitative analysis of the binding affinities of deletion mutants of FinP and *traJ* RNA has identified the 45-nucleotide SLII region of FinP as a minimal, high affinity binding site for FinO (Jerome & Frost, 1999). SLII consists of a 14 base-pair stem, capped with a seven-residue loop, and flanked by 5' and 3' single stranded tails of four and six nucleotides, respectively. Electrophoretic mobility shift assays have demonstrated that FinO binds this RNA as a monomer (Chapter 3). The C-terminal helix of FinO, important in RNA recognition *in vitro*, has also been shown to be important for protecting the nucleotides in the 5' single stranded linker of FinP from RNase E degradation *in vivo* (Sandercock & Frost, 1999). This suggests that the single stranded tails at the base of SLII interact with the positive face of the central domain of FinO (Figure 4.6A). We predict that the SLII stem extends up from the central domain of FinO, along the positively charged N-terminal helix. Interestingly, the length of the N-terminal helix (45 Å) matches the predicted length of the SLII stem-loop. Positioned within the positively charged region near the N-terminus of helix 1 is an exposed tryptophan (Trp-36). It is tempting to speculate that this Trp might stack with unpaired residues in the loop. Similar stacking interactions have been observed between aromatic amino acid residues and unpaired bases in ribonucleoprotein (RNP)–RNA complexes and in a complex between a helical peptide derived from the λN transcriptional antiterminator and an RNA hairpin (Price *et al.*, 1998; Oubridge *et al.*, 1994; Legault *et al.*, 1998).

The electrostatic potential surface of FinO reveals a large negatively charged region, distinct from the proposed RNA binding surface (Figure 4.4A,B). The negatively charged region, which is best viewed at the bottom of the molecule in Figure 4.4B, is composed of residues from helix 3, the β3- β4 hairpin, and the N-terminal end of helix 6. Interestingly, the ColE1 Rop protein dimer that is also involved in facilitating sense–antisense RNA interactions, also has separate positive and negatively charged surfaces (Predki *et al.*, 1995, Banner *et al.*, 1987). Electrostatic repulsions between the negatively charged surfaces of these proteins and their respective RNAs might help to restrain the orientation of the RNA on the protein.

Our model assumes that the conformation of FinO does not change in response to RNA binding. We have shown, however, that helix 1 is flexible and is packed against the remainder of the protein through a modest contact with helix 6 that buries only 408 $Å^2$ of solvent accessible surface area. It is therefore possible that this helix could adopt a different orientation with respect to the rest of the protein in complex with RNA. A similar conformational adjustment in the C-terminal-most helix in the U1A RNP domain is observed upon interaction with RNA (Oubridge *et al.*, 1994). Studies are currently underway to determine the crystal structure of FinO bound to a minimal RNA substrate to directly determine the manner by which FinO contacts RNA.

4.2.5 FinP-*traJ* RNA association

Antisense–sense RNA recognition in natural bacterial systems is usually initiated by base pairing interactions between complementary loops in a 'kissing' complex (Wagner & Simons, 1994). Most of these loops (including those in FinP and *traJ* RNAs) contain a 5'-YUNR-3' motif, first observed in the anticodon loops of tRNAs (Franch *et al.*, 1999). YUNR loops contain stacked bases in the 3' portion of the loop that are pre-aligned for Watson-Crick interactions with a complementary single stranded RNA (as in codon–anticodon recognition) or with other loops. Mutations in the loop regions of FinP and *traJ* RNAs, but not in the stem or single stranded linker regions, cripple their ability to interact and effect repression. This strongly suggests that loop–loop interactions play an important part in FinOP mediated repression (Koraimann *et al.*, 1991).

Our model of FinO in complex with FinP SLII suggests that FinO interacts with RNA to position the N-terminus of helix 1 near the unpaired loop (Figure 4.6A). While the highest affinity-binding site for FinO within FinP is SLII, FinO can also bind to a similar, complementary stem-loop structure in *traJ* RNA. Assuming that FinO interacts with both stem-loop structures in a similar way, we predict that the N-termini of the two FinO molecules bound to the complementary stem-loops will be in close proximity in a FinP–*traJ* RNA kissing complex (Figure 4.6B).

Residues 1–25 of FinO do not play a role in binding individual RNA substrates (Ghetu *et al.*, 1999), but it is possible that this positively charged region could specifically interact with, and stabilize, loop–loop interactions in a kissing complex. To test this prediction, we determined the relative abilities of FinO and FinO(26–186) to promote FinP–*traJ* RNA pairing in an *in vitro* duplexing assay (Figure 4.7A,B; see Experimental Procedures). While FinO dramatically enhances the rate of FinP–*traJ* RNA hybridization, we found that the enhancement is significantly less with FinO(26–186). The apparent rate constants for RNA hybridization, calculated from these and similar experiments, are: 5 (±2) x 10^5 M^{-1} s^{-1} with no protein; 2.5 (±1) x 10^7 M^{-1} s^{-1} with FinO; and 2.9 (±1) x 10^6 M^{-1} s^{-1} with FinO(26-186). Thus, deletion of the N-terminal 25 amino acids of FinO reduces the rate of RNA association ~10-fold compared to the full-length protein.

4.3 Discussion

We have shown that N-terminal regions of FinO specifically enhance the rate of FinP–*traJ* RNA hybridization, perhaps through direct stabilization of a kissing RNA intermediate. Thus, FinO may act in a manner analogous to Rop, which stabilizes the association between the loops of RNAI and RNAII to regulate ColE1 plasmid replication (Eguchi & Tomizawa, 1991). FinO, however, adopts a structure that is unrelated to the compact, four-helix bundle structure of the Rop dimer (Banner *et al.*, 1987). The highly extended structure of FinO is probably necessary to allow this relatively small protein to contact the interacting loops in the kissing complex, as well as single stranded regions in FinP that are otherwise highly susceptible to RNase degradation.

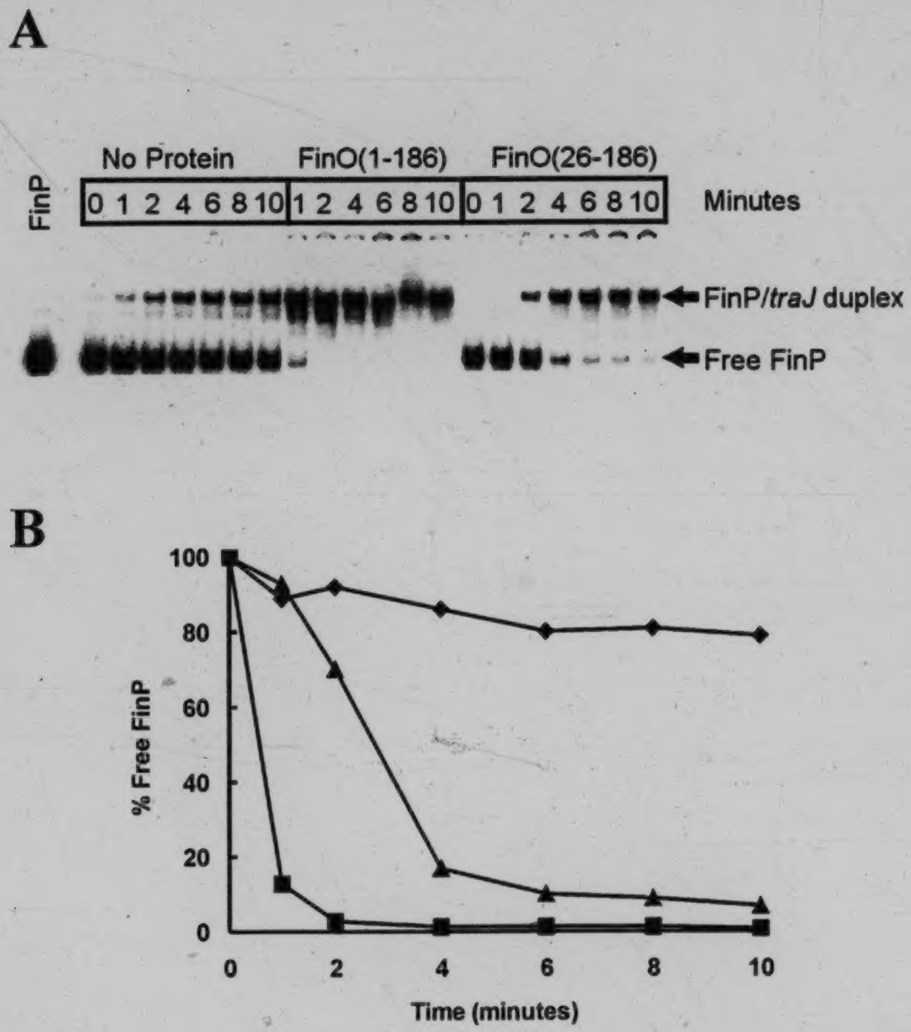


Figure 4-7: **(C)**, Analysis of the rate of duplex formation between 32 P-FinP and 3 H-*traJ m*RNA in the presence and absence of FinO(1–186) or FinO(26–186). In all cases, a 10-fold molar excess of *traJ* RNA was present in the duplex reaction. Duplex formation in the absence of protein was compared to duplex formation in the presence of 0.5 µM FinO(1–186) and 0.4 µM FinO(26–186). The amount of protein added was established by the amount of protein required to bind ~100% of 7.5 fmoles of FinP in a standard gel mobility shift assay. The rate of transfer of free FinP into the duplex with *traJ* RNA was used to calculate the apparent association rate constants. **(D)**, The amount of free FinP RNA remaining at each time point during the duplexing assay shown in **(C)** is plotted for FinO(1–186) (squares), FinO(26–186) (triangles), and no FinO (diamonds). These experiments were performed by Michael Gubbins.

Chapter 5

Probing FinO – FinP RNA interactions by site-directed protein-RNA crosslinking and gelFRET*

---

* A version of this chapter has been submitted for publication in *RNA* (October 2001), Alexandru F. Ghetu, Tom K. Kerppola, and J. N. Mark Glover

5.1 Introduction

Plasmid conjugation is a major mechanism for transfer of antibiotic resistance and virulence determinants among bacteria (Mazodier & Davies, 1991). Perhaps the best studied group of conjugative plasmids is the F family. F plasmids contain a large, multi-cistronic transfer (*tra*) operon that encodes the majority of proteins required for conjugation (Frost et al., 1994). Transcription of the *tra* operon is positively regulated by the plasmid-encoded product of the *traJ* gene (Mullineaux & Willetts, 1985). The production of TraJ is, itself, negatively regulated by a two-component repression system consisting of FinP, an RNA that is antisense to the 5' end of *traJ* mRNA, and FinO, a 21.5 kDa RNA binding protein (Finnegan & Willetts, 1972). Duplex formation between FinP and *traJ* mRNA occludes the ribosomal binding site and prevents translation of *traJ* RNA (van Biesen & Frost, 1994). An RNaseE recognition site located between the two stem-loops of FinP makes this RNA highly susceptible to degradation (Jerome et al., 1999). FinO binds to both FinP and *traJ* mRNA, protecting FinP from degradation and enhancing the rate of duplex formation between FinP and *traJ* mRNA (Chapter 4; Jerome et al., 1999; Koraimann et al., 1996; Lee et al., 1992; van Biesen et al., 1993).

Recent results have begun to reveal the molecular mechanism underlying FinOP function. Biochemical studies have shown that FinO binds as a monomer to stem-loop structures with short 5' and 3' single-stranded tails, such as SLII in FinP and the complementary stem-loop in *traJ* RNA (Chapter 3; Jerome & Frost, 1999) (Figure 5.1). The crystal structure of FinO revealed an elongated, largely helical molecule reminiscent of a right-handed fist with an extended index finger and a thumb that touches the index finger near its base (Figure 5.3). The finger corresponds to a solvent-exposed N-terminal helix (α1), whereas the thumb corresponds to the C-terminal-most helix (α6). An extended, positively charged surface composed of parts of α1, α6 and the fist was suggested to contact RNA targets (Chapter 4) (Figure 5.2). The N-terminal 25 residues of FinO are not present in the crystal structure and are not required for binding to individual RNA substrates, however, these residues facilitate sense-antisense RNA interactions between FinP and *traJ* RNAs (Chapter 4). We have suggested (Chapter 4) that the N-terminal 25 residues of FinO may interact directly with an initial

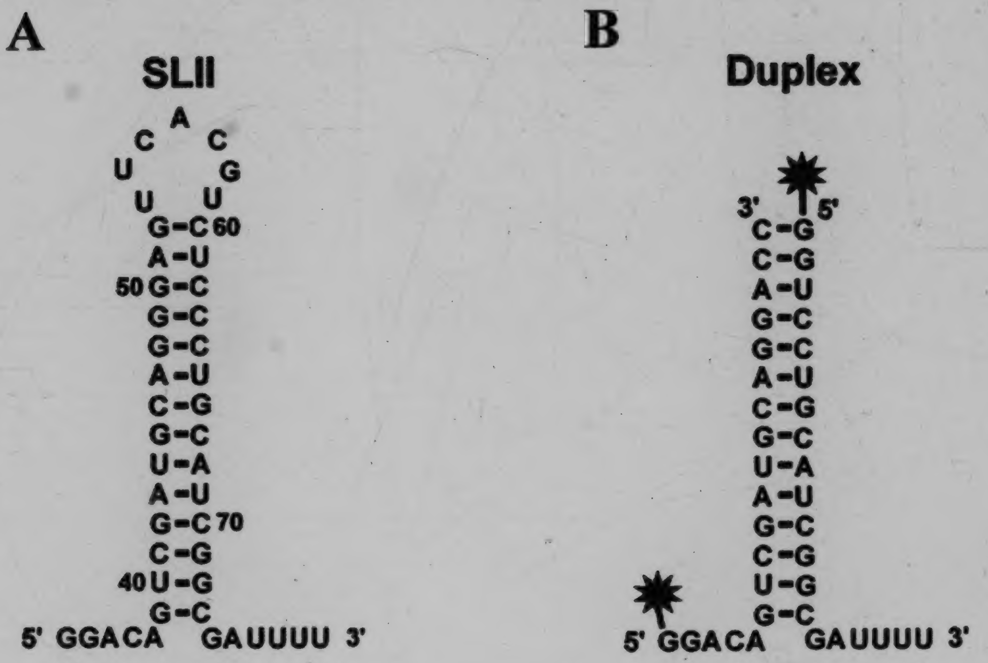


Figure 5.1: Schematic representation of SLII-based RNAs used in this study. **(A)** The nucleotide sequence and predicted secondary structure of SLII. **(B)** The nucleotide sequence and secondary structure of the RNA duplex used in this study. Sites where fluorescein has been attached to the duplex are indicated with stars. The duplex differs from SLII in that the loop is absent and the first three base pairs at the top of the stem are re-ordered.

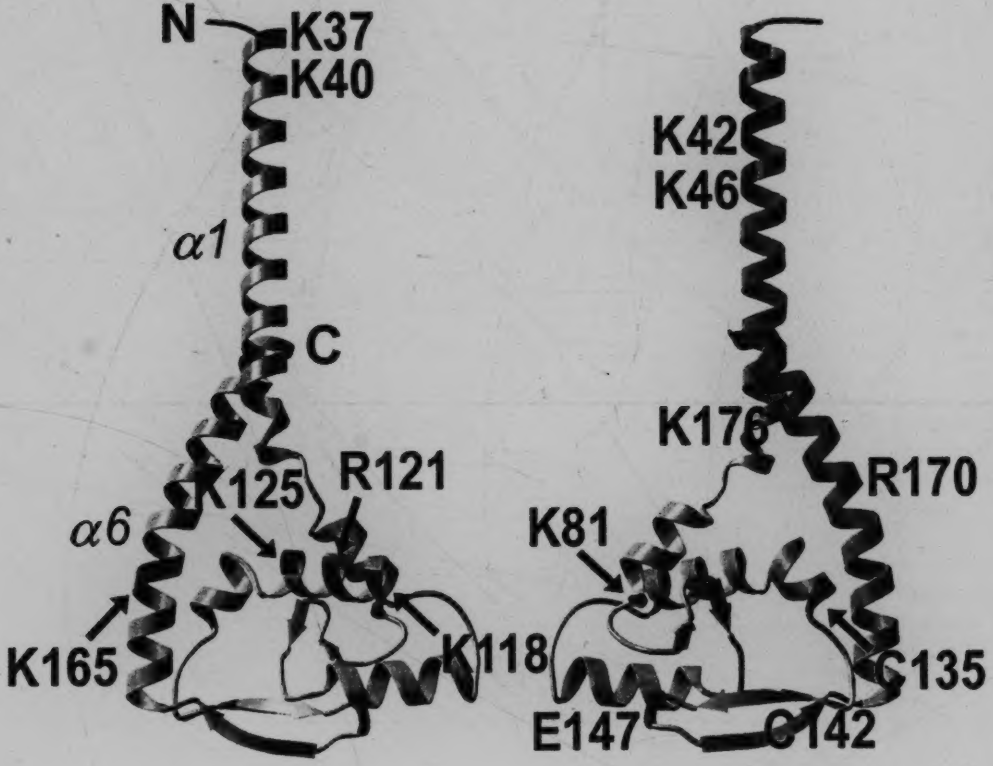


Figure 5.2: Ribbon diagrams of FinO showing the protein in two orientations that are related by an 180° vertical-rotation axis. The molecule on the left-hand side of the panel is considered the front face of FinO (Chapter 4). The positions of cysteine-substituted residues that have been conjugated to APA or Texas Red moieties in the cross-linking and gel-FRET assays are indicated in red.

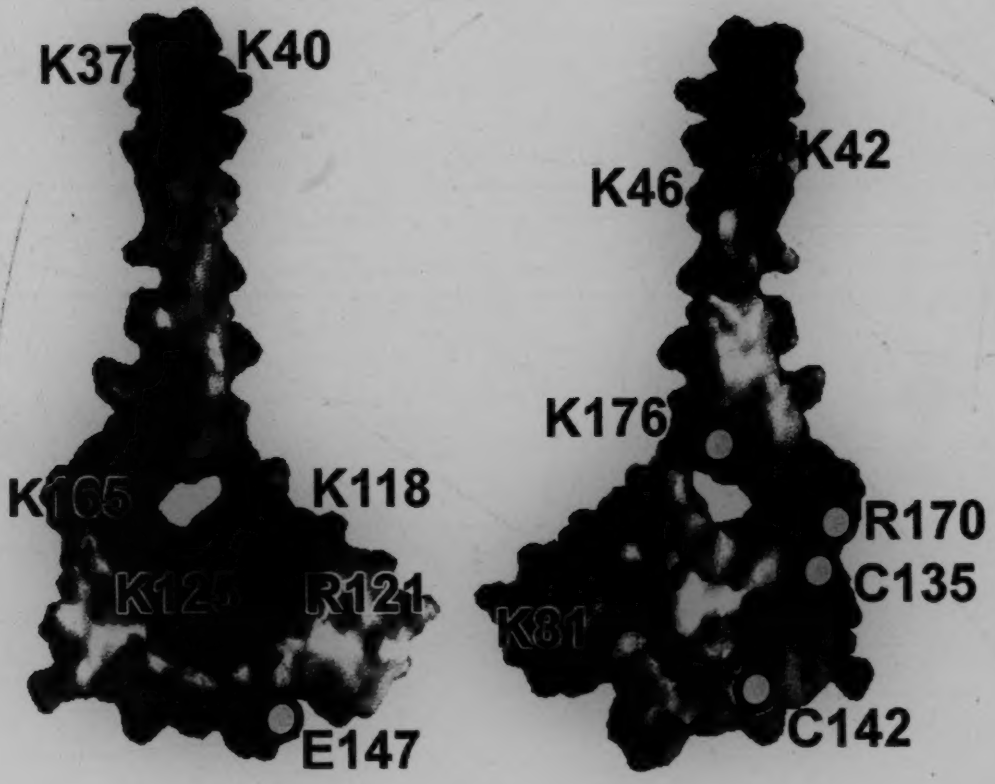


Figure 5.3: Electrostatic surface potential maps of FinO with the protein in the same orientations as in Figure 5.2. Positive regions are indicated in blue, while negative regions are indicated in red. The positions of cysteine-substituted residues that have been conjugated to APA or Texas Red moieties in the cross-linking and gel-FRET assays, respectively, are indicated. Purple spots indicate sites forming significant crosslinks to RNA, while yellow spots indicate sites forming weak or no crosslinks to RNA. Position 118 was not used in the crosslinking assays (see section 5.2.1 for more details).

"kissing complex" formed between complementary loops in FinP and *traJ* RNAs (Koraimann et al., 1991), thereby facilitating FinP-*traJ* RNA recognition. This possibility, together with the structure of FinO, provided the basis for a model of FinO bound to SLII from FinP (Chapter 4). In this model, the stem-loop of SLII lies along the exposed, positively charged α1 of FinO so that the N-terminus is positioned near the SLII loop to participate in loop-loop recognition. The single-stranded tails at the base of the stem interact with a large positively charged patch on the globular body of the protein.

We have used site-specific protein-RNA cross-linking and a gel-based fluorescence resonance energy transfer (gelFRET) assay to investigate the interactions between FinO and SLII. The results of the cross-linking experiments reveal an extensive surface on FinO that comes into contact with RNA. The gelFRET experiments have allowed us to map the proximity of specific sites on FinO and FinP and have indicated that the single-stranded tails at the base of the duplex are in much closer proximity to FinO than is the opposite end of the duplex proximal to the loop. These data suggest that FinO binding to SLII RNA may involve conformational changes in FinO, SLII, or both.

5.2 Results and Discussion

5.2.1 Site-specific FinO-SLII RNA cross-linking

To determine the regions of FinO that are in close proximity to the target RNA, we used a site-specific cross-linking assay involving the photo-activated cross-linker, azidophenacyl bromide (APA-Br). We first replaced the three native cysteine residues in FinO with serines and used a gel electrophoretic mobility shift assay to show that these three substitutions do not alter the affinity of FinO binding to the RNA (Arthur and Glover, unpublished results). We next created a set of 14 FinO mutants with each containing a single cysteine substitution at various solvent-exposed positions, to which we could attach APA via a thioester linkage (Figures 5.2 and 5.3). The sites of substitution include the positively charged surfaces on the tip of the N-terminal helix, the body of the protein, the C-terminal helix and the negatively charged surface on the bottom of the molecule, as shown in figure 5.2.

APA-modified FinO mutants were incubated with SLII RNA, a minimal RNA target for FinO (Jerome & Frost, 1999). Cross-linking was induced by irradiation of the

APA-modified protein/RNA complexes with UV light, thereby activating the azido functional group of APA. The resulting nitrene reacts in a non-specific manner with protein or RNA that is within an ~10 Å radius of the modified cysteine (Chen & Ebright, 1993; Pendergrast et al., 1992). The reaction products were separated by a denaturing polyacrylamide gel and the gel was stained with Coomassie blue and ethidium bromide to visualize protein and RNA, respectively (Figure 5.4). FinO-SLII cross-linking was detected by the presence of high molecular weight species that contain both protein and RNA. As a negative control, the APA-treated cysteine-free FinO mutant protein does not cross-link to SLII under these conditions.

APA-modified residues in the positively charged tip of the N-terminal FinO helix (residues 37, 40, 42 and 46) were all able to cross-link to SLII to a significant extent. In addition, APA-modified residues 121 and 125, located in the positively charged region of the globular body of FinO, also exhibited relatively efficient cross-linking to SLII. These results confirm that these two positively charged regions are in close proximity to the RNA substrate and likely play a significant role in recognizing target stem-loop structures. A small amount of cross-linking was observed when APA was positioned at residues 81, 165 and 176. Residues 165 and 176 are found on the C-terminal helix of FinO (the "thumb") that contains several positively charged residues. Residue 81 is on the body of FinO, but on the opposite face from the major positively charged surface. As predicted (Chapter 4), cross-linking was not observed when APA was positioned on the negatively charged region of FinO (residues 142 and 147). These results are summarized in Figure 5.3 and indicate that the N-terminal helix and the positively charged surface on the body of the protein are in closest contact with RNA. However, interactions also occur on the opposite face of the protein, possibly due to the wrapping of the SLII tails around FinO.

The degree of cross-linking decreases as the APA moiety is placed further down the N-terminal helix of FinO (37>40>42>46). This trend suggests that residues at very N-terminus of α1 are closer to SLII than more C-terminal residues in α1. Ethidium bromide staining revealed the presence of additional higher molecular weight products when the APA was placed on the N-terminal tip of α1 at residues 37, 40 and 42. FinO has been shown to bind SLII non-specifically at high concentrations and in molar excess of the

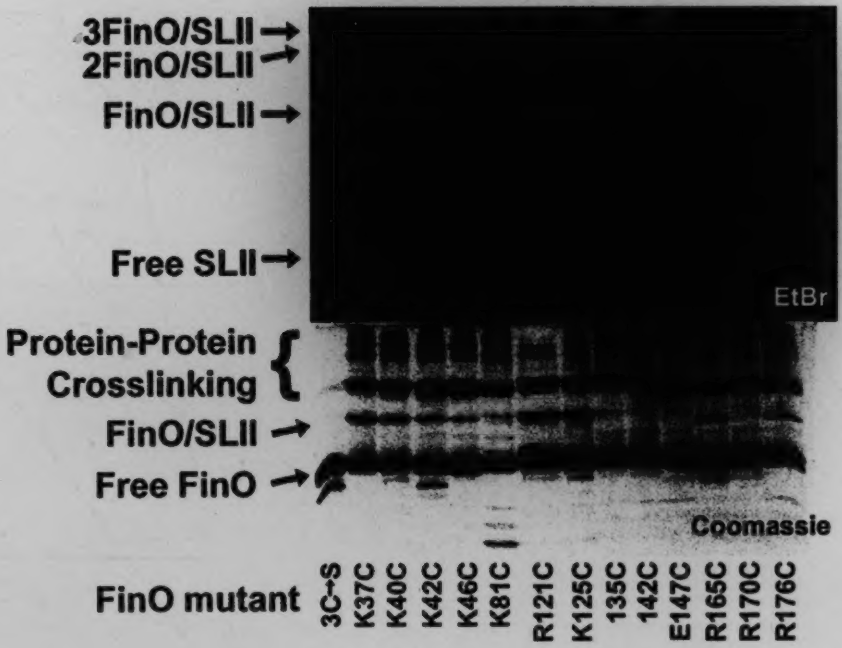


Figure 5.4: Site specific cross-linking of FinO and SLII. Single cysteine FinO mutants conjugated to APA were bound to SLII and cross-linked by illumination at 302 nm. The products of the reactions were separated by 15% denaturing PAGE. RNA was then visualized by ethidium bromide staining (top), while proteins were detected by Coomassie blue staining (bottom). Indicated on the bottom of the figure are the various cysteine mutants used. Indicated on the left of the figure are the position of free SLII, free protein and cross-linked product containing one protein bound to one SLII (resulting from specific interactions) or multiple proteins associated with a single SLII molecule (resulting from specific and non-specific interactions). Additional bands on the Coomassie stained gel are due to non-specific cross-linking occurring between FinO molecules.

RNA target (Jerome & Frost, 1999). For this reason, we believe that these higher molecular weight products are likely due to the binding of two or more FinO proteins to a single SLII molecule. Since these bands are only observed when APA is attached to the N-terminal helix, we suggest that this helix is the main region of FinO involved in non-specific binding to RNA.

5.2.2 Probing FinO-RNA architecture using gelFRET

We have previously suggested that FinO and SLII RNA interact such that the long axis of FinO is parallel to the stem, the positively charged surface on the core of FinO is in contact with the base of the stem, and the N-terminus of $\alpha 1$ lies near the SLII loop. To test this model, we used a gel-based fluorescence resonance energy transfer (gelFRET) analysis (Ramirez-Carrozzi & Kerppola, 2001a). The gelFRET approach is based on separation of the fluorescent complexes by gel electrophoresis and analysis of the separated complexes by scanning of the gel using a laser that excites the donor fluorophore in the gel. In this way, fluorescence from unbound protein, RNA and non-specific protein-RNA complexes can be eliminated and the relative efficiencies of energy transfer between donor and acceptor molecules at different positions on the protein and nucleic acid components can be assessed. This method has been used to determine the orientation of Fos-Jun heterodimer binding at different AP-1 binding sites (Diebold et al., 1998; Leonard & Kerppola, 1998; Ramirez-Carrozzi & Kerppola, 2001b; Ramirez-Carrozzi & Kerppola, 2001c; Ramirez-Carrozzi & Kerppola, 2001d). Here we present the first use of gelFRET to study protein-RNA complexes.

For our experiments, we have prepared two different RNA duplexes that are labeled on their 5' ends with the donor fluorophore fluorescein. The 5' ends of the two strands are located at opposite ends of the duplex RNA (Figure 5.1). These RNAs are based on SLII, but lack the single-stranded loop that connects the two strands in SLII. Using native gel electrophoresis we determined dissociation constants of 4.8 ± 0.3 nM and 4.1 ± 0.5 nM for FinO binding to SLII and the RNA duplex, respectively. These results ensure that FinO binds the duplex substrate in a similar fashion to SLII and confirm previous results that the loop has no significant effect on the interaction between FinO and SLII RNA (Jerome & Frost, 1999). Our binding affinities are ~ 20 times tighter than those determined previously (Jerome & Frost, 1999). These differences probably reflect changes in the gel mobility shift assay and improvements in the purification and quantitation of FinO.

FinO mutants labeled with Texas Red were mixed with an equi-molar amount of RNA labeled with fluorescein. The nucleoprotein complexes were separated from unbound FinO, unbound duplex and free Texas Red by non-denaturing gel electrophoresis. The gel was scanned using a 488 nm argon-ion laser that excites fluorescein. The fluorescence emissions of both the donor fluorescein and acceptor Texas Red were measured at each position in the gel. Labeling of FinO by the Texas Red fluorophore altered the gel electrophoretic mobility for most of the FinO-RNA complexes (compare the mobilities of complexes labeled with Texas Red (+) to unlabeled complexes (-) in figure 5.5). In cases where the complexes with and without acceptor could be separated, it was possible to ascertain that the analysis of the nucleoprotein complexes would not be influenced by the presence of unlabeled protein. The percentage of protein modified with Texas Red depended on the location of the cysteine residue, but was calculated to be greater than 60% in all cases examined (see Experimental Procedures). The relative efficiencies of energy transfer were determined by comparing the ratios of acceptor to donor fluorescence for each combination of acceptor and donor fluorophores.

In general, energy transfer between the RNA and each of the labeled FinO proteins was higher when the fluorescein was positioned on the 5' single-stranded tail of the RNA (left most lanes, Figure 5.6) compared to the opposite end of the duplex stem, where the loop would be found in SLII (right most lanes, Figure 5.6). In contrast, nonspecific complexes formed by proteins labeled at residues 125 and 176 exhibited significant FRET when the fluorescein was positioned at the opposite end of the RNA duplex, indicating that fluorescein linked to this position is capable of efficient energy transfer to Texas Red (Figure 5.6). Thus, the fact that efficient energy transfer in specific FinO-RNA complexes is only observed when the RNA is labeled on the 5' single stranded tail indicates that FinO binds closer to the unpaired single strands than to the distal end of the duplex stem.

In chapter 3, we suggest that the tip of the N-terminal helix of FinO contacts the RNA in or near the loop of SLII. However, our gelFRET results indicate that this part of

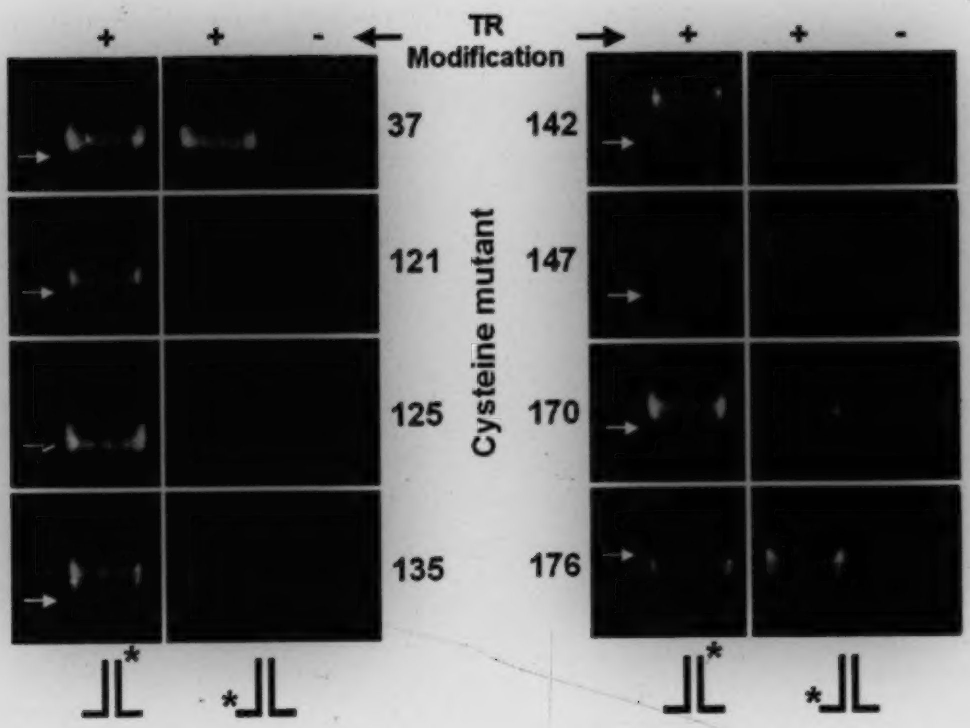


Figure 5.5: Labeling of FinO with Texas Red leads to changes in the mobilities of several of the FinO-SLII complexes. Shown in this panel are the band-shifts of Texas Red modified (lanes 1 and 2) or unmodified (lane 3) proteins in complex with duplex, labeled with fluorescein either on the 5' tail (lane 1) or at the top of the stem (lanes 2 and 3). Each row represents a different cysteine point mutant, as indicated in the middle of the figure. Red arrows indicate complexes with modified protein, while white arrows indicate complexes with unmodified proteins. Changes in mobility varied from one mutant to another. Although the addition of TR leads to changes in the mobility of most mutants, this did not hold true for K125C.

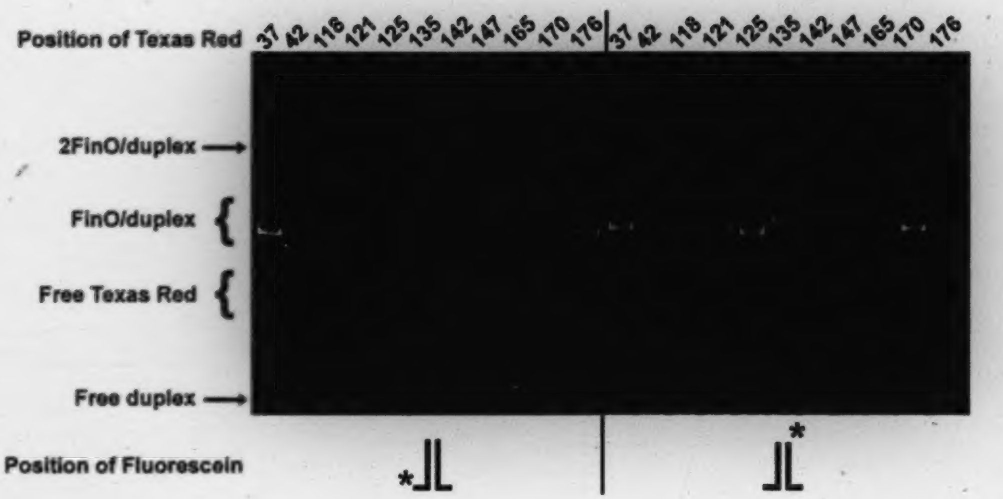


Figure 5.6: Probing the interactions of FinO bound to a target RNA by gelFRET. FinO cysteine point mutants labeled with Texas Red were bound to the RNA duplex containing fluorescein at one of two positions. Complexes were separated from free RNA, free Texas Red and non-specific complexes by 10% non-denaturing PAGE. The gel was then scanned using a 488 nm argon ion laser to excite the fluorescein on the RNA, and the emission from both the fluorescein donor (green) and the Texas Red acceptor (red) were measured separately to obtain two images that were then superimposed. In the bands corresponding to the FinO-SLII complex, the color reflects the distance between the fluorophores, with a red color indicating closer distances and higher energy transfer than a green color. RNA duplex was labeled with fluorescein either on the 5' tail (left hand lanes) or the top of the stem (right hand lanes). Indicated on the left of the figure are the positions of the free duplex RNA, FinO-duplex, non-specific complexes and free un-incorporated Texas Red. The cysteine substitutions used are indicated above the lanes.

the N-terminal α-helix is, instead, much closer to the 5' single-stranded tail. These data suggest that it is likely that either FinO or the RNA undergoes a conformational change upon complex formation. It is possible that the solvent exposed N-terminal helix of FinO might rearrange to allow its positively charged N-terminus to come into closer proximity with the main body of the protein, when bound to RNA. Indeed, the N-terminal helix is more flexible than the globular core of the protein, as shown by its susceptibility to limited proteolysis (Chapter 3) and its overall high crystallographic B factor, relative to the rest of the structure (Chapter 4).

Interestingly, we have recently discovered that FinO can unwind its bound RNA substrate (discussed in chapter 6). Because the observed energy transfer is a weighted average of the energy transfer of the individual conformational states present during the measurement, such dynamic conformational changes in RNA structure could also have complex and profound effects on the observed energy transfer. For example, FinO might make intimate contact with the end of the duplex closest to the unpaired loop, but if the lifetime of this conformational state is short, it will not make a significant contribution to the observed FRET signal. While we hope to probe the static structure of FinO-RNA complexes by X-ray crystallography, it seems likely that an understanding of the dynamic processes that may be critical to FinO function may require other, solution based approaches, such as quantitative gelFRET experiments and NMR.

Chapter 6

The FinO repressor of bacterial conjugation is an RNA chaperone that facilitates sense-antisense RNA interactions*

---

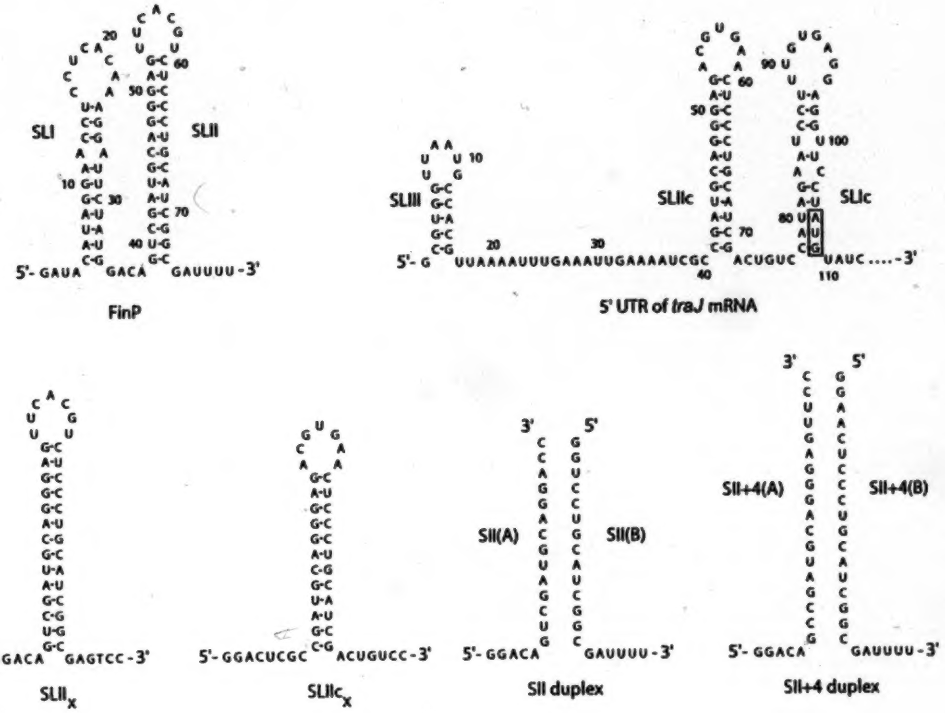
* A version of this chapter is currently under review in Molecular Cell (January 2002), Alexandru F. Ghetu, Michael J. Gubbins, David Arthur, Ross Edwards, Laura S. Frost, and J. N. Mark Glover.

6.1 Introduction

The F family of plasmids confer antibiotic resistance and virulence to a wide variety of enterobacteria. The rapid transfer of F-like plasmids among bacterial species has been linked to the rapid acquisition of antibiotic resistance in strains of *E. coli* that caused wide-spread outbreaks of antibiotic-resistant dysentery in post-World War II Japan (Watanabe & Fukasawa, 1961). More recently transfer has been shown to have occurred between *E. coli* and *Salmonella*, leading to the introduction of virulence operons and pathogenicity in the *Salmonella* type I strains (Boyd & Hartl 1997; Boyd & Hartl 1998).

The RNA binding protein FinO, along with the 79-nucleotide antisense RNA FinP, make up a two-component inhibition system for F-plasmid mediated bacterial conjugation (Finnegan D. & Willetts N. 1972) (Figures 6.1 and 6.2). FinP is complementary to the 5' untranslated region of *traJ* mRNA and blocks ribosomal entry when associated with this mRNA (Mullineaux & Willetts 1985; van Biesen & Frost 1994). TraJ is a transcriptional activator that is required for expression of the majority of conjugative protein components (Cuozzo & Silverman 1986). In the absence of FinO, FinP is rapidly degraded by RNases within bacterial cells, allowing TraJ to be readily synthesized (Lee *et al.*, 1992). FinO binds FinP and *traJ* mRNA, stabilizing FinP against degradation and promoting extended duplex formation between the complementary RNA molecules (Lee *et al.*, 1992; Jerome *et al.*, 1999; van Biesen & Frost, 1994). This, in turn, blocks TraJ translation and inhibits bacterial conjugation.

Recent biochemical and crystallographic studies have begun to reveal how FinO interacts with its target RNAs. FinO recognizes RNA duplex stems that have 5' and 3' single stranded tails at one end (Jerome & Frost, 1999). As this recognition is dependent on the structure but not the sequence of the RNA, FinO can bind several such tailed stem-loop structures in both FinP and *traJ* RNAs (Jerome & Frost, 1999). FinO adopts an elongated structure with a solvent exposed N-terminal helix extended from a C-terminal domain (Chapter 4). Biochemical studies have demonstrated that FinO binds stem-loop structures as a monomer and that positively charged surfaces in the N-terminal helix and the globular body of the protein directly contact RNA (Chapters 3 and 5).



FinP

5' UTR of *traJ* mRNA

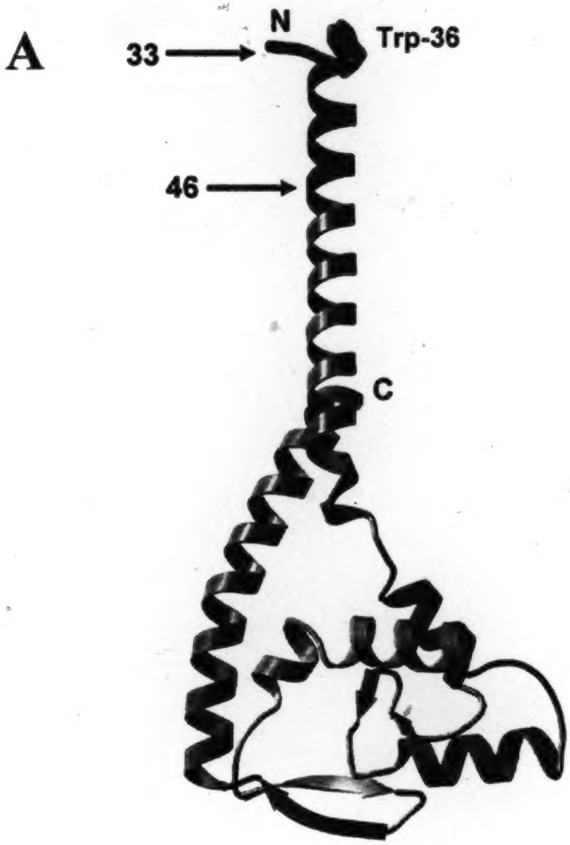
$SLII_x$

$SLIIc_x$

SII duplex

SII+4 duplex

Figure 6.1: The sequence and secondary structure of RNA molecules used in this study. Only the 5' UTR of *traJ* mRNA (which duplexes with FinP) is shown and the start codon is boxed. The A and B strands of SII and SII+4 are aligned to show the regions of base-pair complementarity.



B

TIINVTTPPKWKVKKQKLAEL

30 35 40 45

Figure 6.2: FinO tertiary fold and N-terminal sequence. (**A**) Ribbons representation of FinO, with Trp-36 displayed in yellow. (**B**) Amino acid sequence of the N-terminal region of FinO that is critical to its ability to unwind duplex RNA.

Pairing of FinP and *traJ* RNAs is believed to initiate with the formation of "kissing complexes" between complementary loops in the two RNAs (Koraimann *et al.*, 1991; Finlay *et al.*, 1986) and it has been suggested that FinO stabilizes kissing complexes to facilitate sense-antisense RNA interactions (Chapter 4). The 25 N-terminal amino acids of FinO enhance the rate of FinO mediated FinP-*traJ* duplexing 10-fold, but do not play a significant role in the binding of individual RNA targets (Chapters 3 and 4). Based on these observations, it has been suggested that, upon binding stem-loop structures, the N-terminal region of FinO is positioned near the RNA loop to stabilize loop-loop pairing. A similar function has been demonstrated for Rom, the protein that binds and stabilizes sense-antisense RNA kissing complexes to ultimately inhibit replication of the ColE1 plasmid (Eguchi & Tomizawa, 1991; Predki *et al.*, 1995). Interestingly, both FinP and *traJ* mRNA contain stem-loop structures that are predicted to be very stable. These stems would be expected to present a kinetic barrier to duplex formation and it has been previously suggested that FinO might act to destabilize intramolecular base-pairing within its RNA target to allow the formation of sense-antisense interactions (Chapter 3).

Unwinding of double stranded DNA and RNA is an essential component of many cellular events including replication, transcription, translation, recombination, DNA repair and nuclear pre-mRNA splicing (for a review, see von Hippel & Delagoutte, 2001). Proteins have been shown to participate in duplex unwinding by either of two general mechanisms. In the first, DNA or RNA helicases hydrolyse ATP to provide the free energy required for base-pair disruption (von Hippel & Delagoutte, 2001; Tanner & Linder, 2001). The second mechanism involves single strand-specific nucleic acid binding proteins (for example, bacterial SSB) (Lohman & Ferrari, 1994), the replication protein A (RPA) (Iftode *et al.*, 1999), as well as RNA-binding proteins such as hnRNP A1 (Herchlag, 1995) and hnRNP C (Shahied *et al.*, 2001), that alter the duplex to single-strand equilibrium by specifically binding and stabilizing the single stranded conformations).

In this study, we show that FinO is able to unwind RNA duplexes by a novel mechanism that does not involve ATP hydrolysis nor the preferential binding and stabilization of single-stranded RNA. Instead, FinO uses its own RNA binding free energy to destabilize a limited number of intramolecular base pairs in its bound substrate. The destabilization of intramolecular base pairing facilitates the formation of intermolecular base pairing between FinO and *traJ* mRNA and is essential for FinO-mediated repression of conjugative plasmid transfer *in vivo*.

6.2 Results

6.2.1 FinO can unwind dsRNA

We have previously suggested that FinO might facilitate sense-antisense RNA interactions by binding to kissing complex intermediates that are proposed to proceed to a full sense-antisense duplex (Chapter 4). However, both FinP and *traJ* RNAs contain large intramolecular duplex regions that are predicted to be very stable. For example, the free energy of unfolding at 37 °C for the SLI and SLII stem-loop structures in FinP are predicted to be -10 and –28 kcal/mole, respectively (Mathews *et al.*, 1999). We therefore wondered if FinO might overcome kinetic barriers to sense-antisense RNA interactions through the specific destabilization of intramolecular secondary structures within the target RNAs. To test this idea, we used an RNA unwinding assay, based on previous methods used to characterize ATP-dependent helicases (Jankowskey *et al.*, 2000) (Figure 6.3). In these assays, we used an RNA duplex (SII) that mimics the structure of FinP SLII (Figure 6.1), and binds with similar affinity to SLII. One strand of SII was labeled with ^{32}P and strand dissociation in the presence or absence FinO was monitored by gel electrophoresis. To prevent the ^{32}P-labeled strand from re-associating with the complementary strand after release from the duplex, a large molar excess of an unlabeled version of the ^{32}P-labeled strand was added at the initiation of the reaction.

The results of these experiments show that FinO is capable of unwinding SII RNA in a time-dependent manner (Figure 6.4A,B). In addition, the unwinding activity of FinO was observed to be highly dependent on FinO concentration (Figure 6.5). The dramatic increase in unwinding activity between 0.5 and 1 μM suggests that multiple FinO molecules might cooperate to unwind RNA. However, as the concentration of FinO is further increased, RNA unwinding gradually decreases. The reduction of unwinding activity at the higher concentrations may be due to additional, non-specific, interactions between FinO and RNA. Indeed, gel electrophoretic mobility shift assays

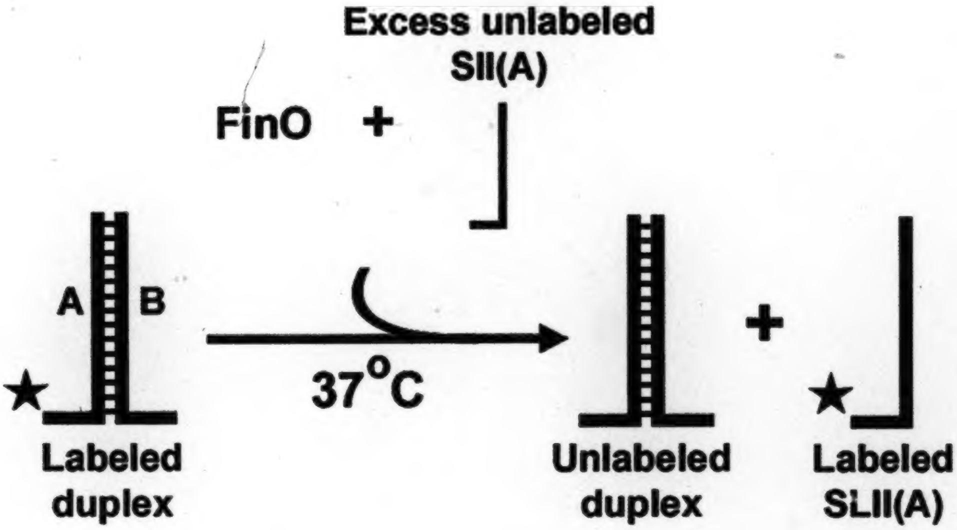


Figure 6.3: Schematic diagram of the unwinding assay.

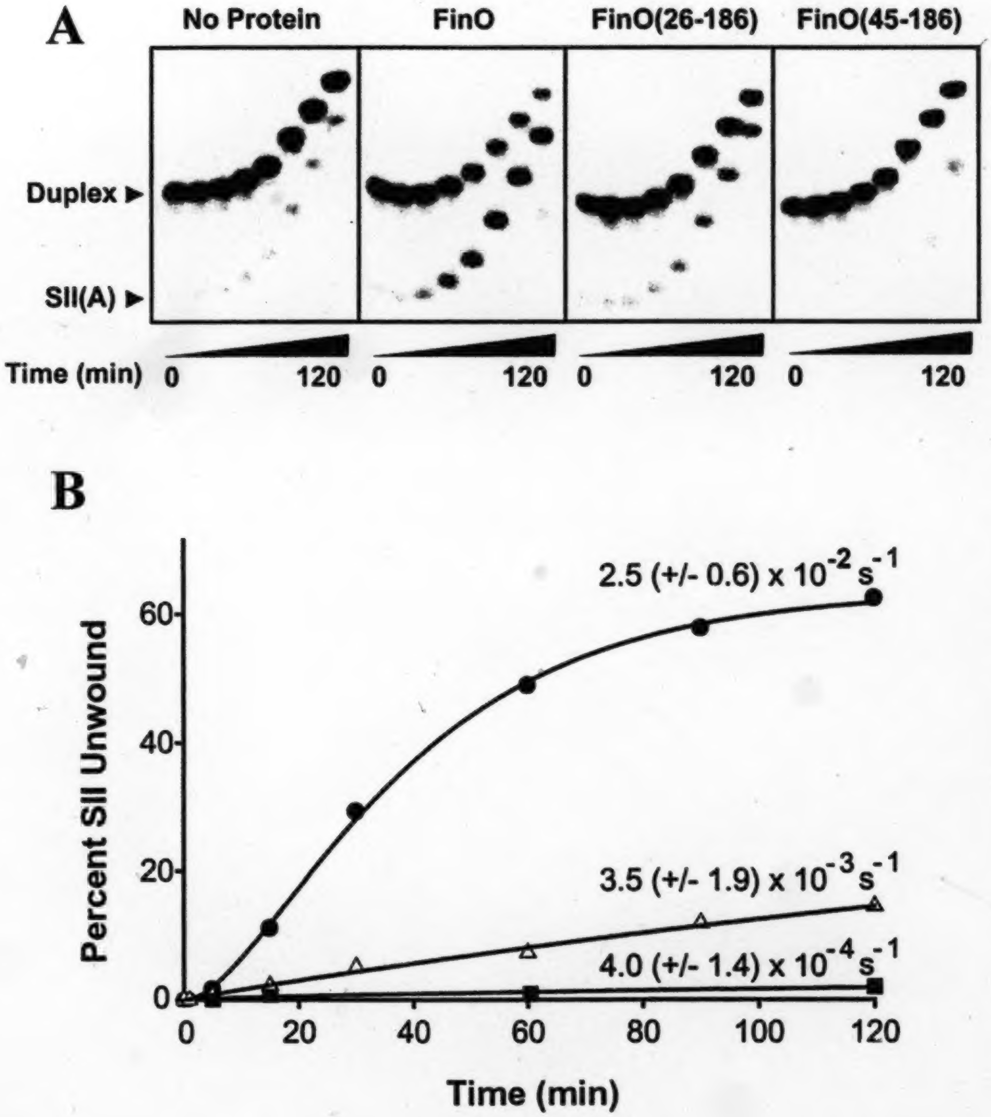


Figure 6.4: N-terminus of FinO is required for unwinding activity. **(A)** Comparison of unwinding efficiencies between FinO and N-terminally truncated fragments. SII was incubated with either FinO, FinO(26-186), or FinO(45-186). Aliquots were taken at 0, 1, 5, 15, 30, 60, 90, and 120 minutes after the start of the reaction and loaded directly onto a continuously running gel. **(B)** The percentage of ^{32}P-labeled SII(A) strand released from the duplex in the unwinding reactions was plotted as a function of time and the apparent first order rate constant, k_1 was determined from this plot (see Experimental Procedures). This work was performed by David Arthur.

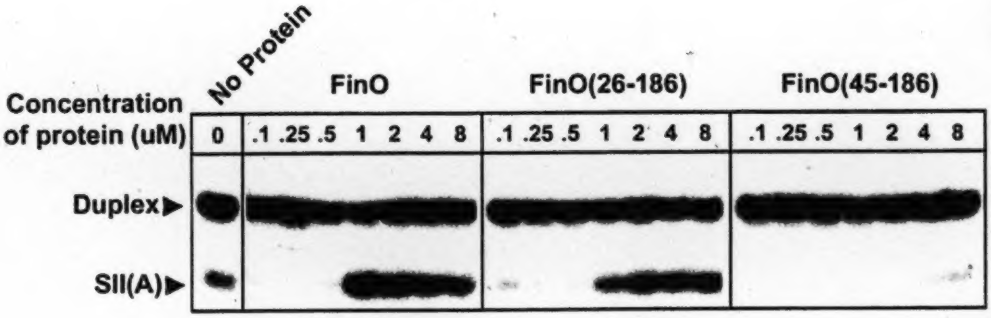


Figure 6.5: Unwinding activity is highly dependent on protein concentration. Unwinding of SII in the presence of either FinO or N-terminally truncated FinO fragments at the concentrations indicated. This work was performed by David Arthur.

indicate that FinO can form large aggregates on RNA at similar concentrations (see Figure 5.5). As expected, factors that stabilize double stranded RNA, such as reductions in temperature or the addition of Mg^{++}, reduced unwinding (Arthur and Glover, unpublished results).

ATP-dependent RNA helicases can unwind large tracts of duplex RNA. To test the processivity of FinO-catalysed unwinding, we assayed the unwinding of an SII-based RNA with a duplex extended by 4 base pairs (SII+4). This RNA is not unwound by FinO as efficiently as SII, indicating that FinO, unlike the ATP-dependent helicases, cannot unwind RNA in a processive manner (Arthur and Glover, unpublished results).

6.2.2 RNA unwinding activity is associated with the N-terminal region of FinO

We previously showed that the N-terminal 25 residues of FinO are dispensable for specific interactions with a single RNA target but nevertheless play an important role in the protein's ability to promote duplex formation between FinP and *traJ* mRNA (Chapter 4). To test the role of N-terminal regions of FinO in RNA unwinding, we compared the abilities of full length FinO and the N-terminal deletion mutants, FinO(26-186) and FinO(45-186), to unwind SII RNA over a two hour time course (Figure 6.4). These data were used to calculate apparent first order unwinding rate constants for each of the proteins (see Experimental Procedures). While FinO(26-186) is able to unwind SII, it does so at about only 1/10 the rate observed for full length FinO at 1 μM protein concentration (Figure 6.4). In contrast, FinO(45-186) appears to be completely deficient in RNA unwinding activity at all concentrations tested, and may even stabilize the duplex form of SII. These results show that the N-terminus of FinO is essential for RNA unwinding activity, and suggest that the region between residues 26 and 44 is particularly critical.

While residues 1-32 are apparently unstructured in the free protein, residues 33-44 constitute the end of a solvent exposed helix that directly contacts RNA (Chapter 4). To test the role of this helix in more detail, we constructed a series FinO mutant proteins with double amino acid-alanine substitutions throughout the region from residue 26 to 44 and measured the abilities of these mutants to unwind SII after a two hour incubation. The results of these experiments (Figure 6.6) show that the extreme N-terminus of the

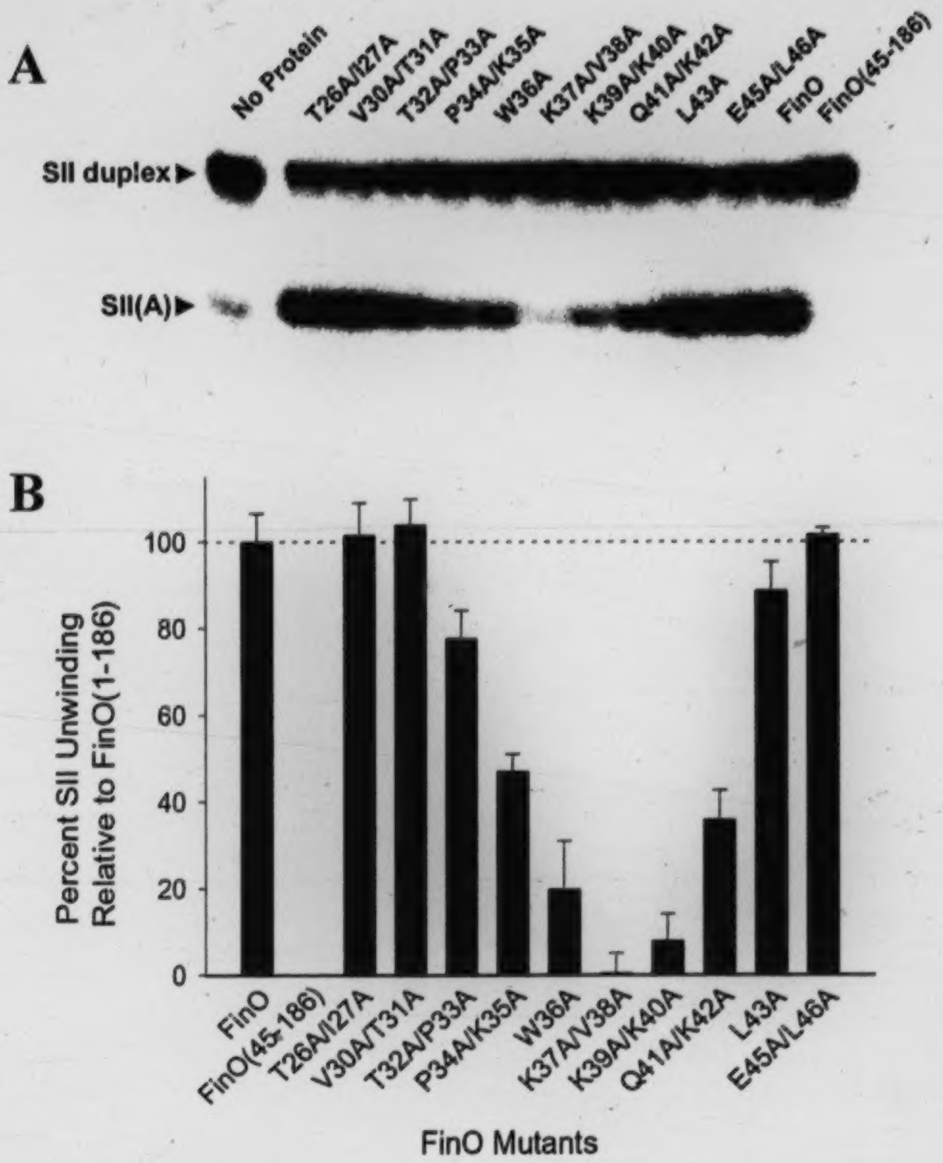


Figure 6.6: Unwinding efficiencies of N-terminal double alanine point mutants. **(A)** Samples were incubated with the indicated double alanine point mutants and unwinding was detected by the release of ^{32}P-labeled SII(A) strand from the SII-duplex. **(B)** The percentage of A strand released from duplex is presented in graphical form. As controls, unwinding of SII in the presence of FinO, FinO(45-186) or no protein was performed in parallel. These experiments were performed with David Arthur.

helix is most important for RNA unwinding. Strand separation of SII wasmost dramatically reduced for the K37A/V38A and K39A/K40A mutants, with less than 30% of the RNA being unwound relative to native FinO. The mutants T32A/P33A, P34A/K35A and Q41A/K42A also had decreased unwinding activity compared to native FinO. To identify the single residues most critical for unwinding, we constructed a second set of single amino acid to alanine substitution mutants and tested their abilities to unwind SII (Figure 6.7). Of the single site substitution mutants, P34A, K35A, W36A, and K40A showed a significant decrease in RNA unwinding activity, with W36A showing the most dramatic effect. Interestingly, the K37A, V38A, K39A and K40A mutants displayed only a slight decrease in SII unwinding over native FinO, whereas the corresponding double point mutants K37A/V38A and K39A/K40A had almost no unwinding activity.

6.2.3 FinO utilizes its RNA-binding energy to unwind RNA

Many DNA and RNA helicases hydrolyze ATP to provide the free energy needed to destabilize base pairing. FinO bears no overall structural similarity to known ATP-dependent helicases and its unwinding activity is not dependent on nucleoside triphosphates (Arthur & Glover, unpublished results). We wondered whether FinO might instead use its free energy of RNA binding to unwind RNA base pairs. To test this idea, we compared the RNA binding affinities of FinO and the unwinding-deficient mutants using an electrophoretic mobility shift assay (Figure 6.8). The results show that the unwinding-deficient mutants all bind SLII RNA with high affinity. Furthermore, we observed an inverse correlation between the RNA binding and unwinding activities of the set of mutants such that the mutant with the lowest degree of unwinding, FinO(45-186), displayed a 20-fold enhanced affinity for SLII over wild type FinO (Table 6.1). We suggest that the overall RNA binding energy is the sum of favorable energy terms derived from protein-RNA contacts, and unfavorable terms derived from the destabilization of base pairing. Removal of the N-terminal regions that are responsible for base pair destabilization reduces the unfavorable energy terms, while leaving most of the favorable interactions intact, resulting in a more favorable overall free energy of RNA binding.

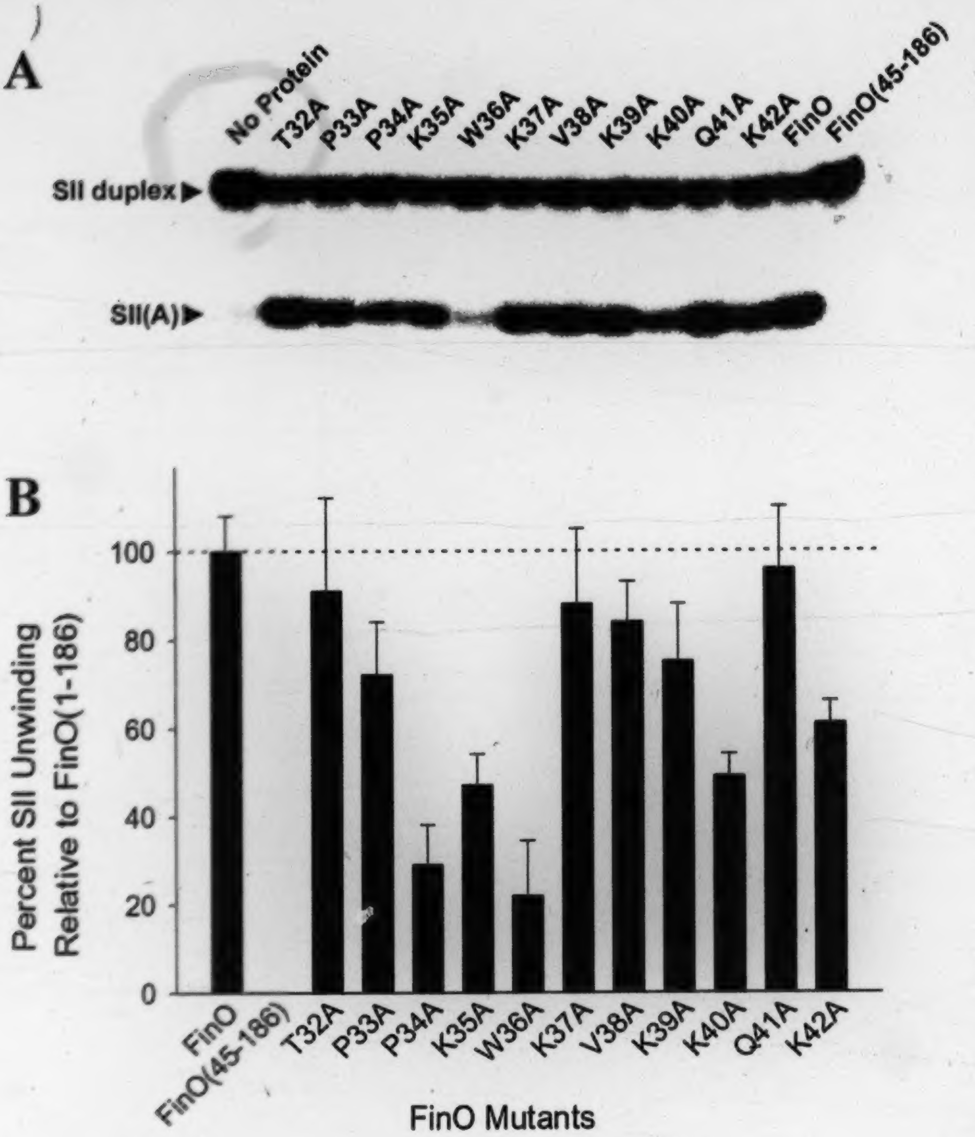


Figure 6.7: Unwinding efficiency of N-terminal single alanine point mutants. **(A)** Samples were incubated with the indicated single alanine point mutants and unwinding was detected by the release of ^{32}P-labeled SII(A) strand from the SII-duplex. **(B)** The percentage of A strand released from duplex is presented in graphical form. As controls, unwinding of SII in the presence of FinO, FinO(45-186) or no protein was performed in parallel. This work was performed by David Arthur.

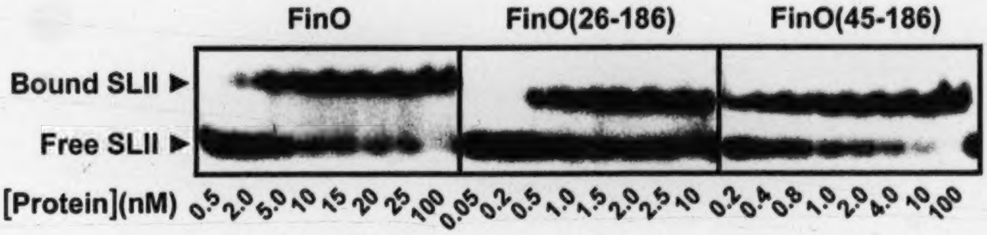


Figure 6.8: Representative EMSAs for FinO, FinO(26-186) or FinO(45-186) binding to SLII. Samples containing 50 nM of SLII were incubated with the protein and the amount of protein indicated. These and similar experiments were used to determine the FinO-RNA dissociation constants shown in Table 6.1.

6.1: Unwinding rates, duplexing rates and mating efficiencies for FinO and FinO derivatives

Protein	Relative affinities[1] (K_a)	Relative rate of unwinding[1] (k_1)	Relative rate of duplexing[1,4] (k_2)		Relative mating efficiency[2,4]
	RNA substrate				
	SLII	SII	SLII/SLIIc	FinP/*traJ*-mRNA	
None		0.02 (± 0.008)	<0.02	0.02 (± 0.008)	1
FinO[3]	1 (± 0.1)	1 (± 0.2)	1 (± 0.4)	1 (± 0.4)	0.05 (± 0.03)
26-186	4 (± 0.3)	0.14 (± 0.08)	0.11 (± 0.07)	0.12 (± 0.08)	0.6 (± 0.4)
45-186	20 (± 2.4)	<0.02	<0.02	0.02 (± 0.001)	1 (± 0.2)
W36A	5 (± 0.9)			0.48 (± 0.2)	0.36 (± 0.1)
P34A	1.67(± 0.1)				0.03 (± 0.02)
K35A	1.14 (± 0.2)				0.03 (± 0.02)
K39A	1.0 (± 0.1)				0.03 (± 0.02)
K40A	0.91 (± 0.1)				0.03 (± 0.02)
K37A/V38A				0.6 (± 0.3)	0.06 (± 0.01)
K39A/K40A	1.37 (± 0.2)			0.4 (± 0.2)	0.02 (0.003)

[1] All rates are as a percentage of FinO

[2] Efficiencies are as a percentage of mating in the absence of protein

[3] For FinO; $Ka = 5\ (\pm 1) \times 10^{7}\ M^{-1}$; $k_1 = 2.5\ (\pm 0.6) \times 10^{-2}$; k_2 (SII_x and $SIIc_x$) $= 1.4\ (\pm 0.2) \times 10^{5}\ s^{-1}M^{-1}$
k_2 (FinP-traJ mRNA) $= 2.5\ (\pm 1) \times 10^{7}\ M^{-1}\ s^{-1}$

[4] Rates of FinP/traJ mRNA duplexing as well as mating efficiencies were determined by Mike Gubbins

6.2.4 Residues 26 to 44 of FinO function to promote FinP-*traJ* RNA duplex formation

To test our proposal that FinO-catalysed RNA unwinding reduces kinetic barriers to sense-antisense duplex formation, we probed the ability of unwinding-deficient FinO mutants to facilitate sense-antisense RNA recognition. We used an *in vitro* RNA duplexing assay to measure the rate of binding of ^{32}P-labelled FinP to its complementary sequence within *traJ* mRNA either alone or in the presence of FinO (Figure 6.9, and Table 6.1). Consistent with previous findings (Chapter 4; van Biesen & Frost, 1994; Sandercock & Frost, 1998), wild type FinO enhances the rate of FinP-*traJ* RNA duplexing 50-fold compared to the no protein control. Second order rate constants (k_2) calculated from these data yield values of 2.5×10^7 M^{-1} s^{-1} and 5.0×10^5 M^{-1} s^{-1} in the presence and absence of FinO, respectively. FinO(26-186) is able to facilitate RNA duplexing, albeit at a 10-fold reduced rate ($k_2 = 2.9 \times 10^6$ M^{-1} s^{-1}) compared to wild type FinO. Duplexing reactions performed in the presence of FinO(45-186) were indistinguishable from no protein controls.

We also analyzed the ability of FinO to promote duplex formation between two complementary stem-loops ($SLII_x$ and $SLIIc_x$) derived from SLII of FinP and SLIIc of *traJ*-mRNA (Figure 6.1). The second order rate constant for duplex formation between the stem-loops was calculated to be 1.4×10^5 M^{-1} s^{-1} in the presence of FinO, while no detectable association occurred in the absence of protein (Figure 6.9, Table 6.1). FinO can therefore promote duplex formation between minimal stem-loop targets and does not require full length FinP and *traJ* mRNA. Consistent with the results obtained with FinP and *traJ* RNA, the rate of $SLII_x/SLIIc_x$ duplex formation in the presence of FinO(26-186) was an order of magnitude lower than with full length FinO and there was no duplex formation in the presence of FinO(45-186). Thus, the N-terminal 44 residues of FinO, which are required for its unwinding activity, are essential to promote the association of FinP and *traJ* mRNA, as well as minimal stem-loop targets.

To assess further the correlation between unwinding and RNA duplexing, we tested the ability of the unwinding deficient FinO double-point mutants to promote duplex formation between $SLII_x$ and a $SLIIc_x$ under the same conditions used in the unwinding assays. We found that there is a significant decrease in the duplexing activity

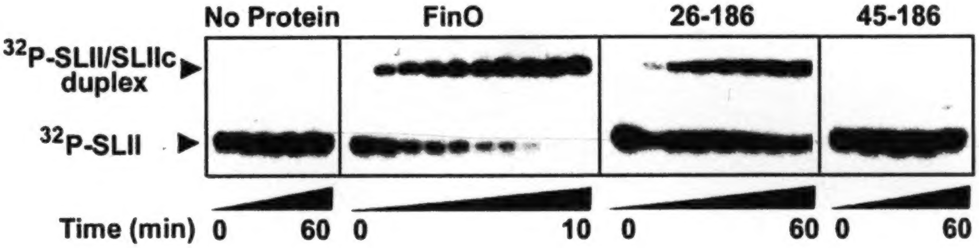


Figure 6.9: Duplexing assays with FinO, FinO(26-186) or FinO(45-186). Proteins were tested for their ability to facilitate sense-antisense pairing between $SLII_x$ and $SLIIc_x$ RNAs and these results were compared with a no protein control. Aliquots were taken at 0, 0.25, 0.5, 0.75, 1, 1.5, 2, 3, 4 and 10 minutes for FinO, 1, 2, 4, 8, 15, 30 and 60 minutes for FinO(26-186) and 0, 10, 20, 30 and 60 minutes for FinO(45-186) and no protein.

of several double-point mutants in comparison to wild-type FinO (Figure 6.10). Alanine substitutions between residues 34 and 42 caused the greatest reduction in duplexing activity, demonstrating that the same residues that are required for efficient RNA unwinding are also critical for RNA duplexing. We conclude that the RNA unwinding function of FinO is mechanistically linked to its duplexing function.

6.2.5 The RNA unwinding and duplexing activities of FinO are essential for the repression of bacterial conjugation

To determine the physiological relevance of the RNA unwinding and duplexing activities of FinO, we assayed the ability of N-terminal mutants of FinO to repress the transfer of F plasmids from donor to recipient *E. coli* cells. The donor cells contained an F-derived plasmid, pOX38, bearing a kanamycin resistance gene, as well as a FinO expression plasmid. pOX38 does not express FinO and is dependent on FinO supplied in *trans* for efficient inhibition of transfer. Donor and spectinomycin-resistant recipient cells were mixed, and transconjugant recipient cells were selected for resistance to both kanamycin and spectinomycin. Mating efficiencies were calculated as the ratio of transconjugants to donor cells.

As expected, mating was severely inhibited by FinO, with only 2% mating efficiency compared to no protein (Table 6.1). FinO(26-186) was a less effective repressor of conjugation and displayed 60% of the mating efficiency seen with no protein, while FinO(45-186) did not inhibit mating at all. These results reveal that the N-terminal region of FinO that contains the RNA unwinding activity plays an essential role in the ability of FinO to block bacterial conjugation.

The alanine substitution mutants used in the unwinding assays were also tested in the mating assay. Of all the mutants tested, only W36A showed a significant loss of repression. The W36A mutant also showed the most dramatic loss of unwinding activity of all the single amino acid substitutions tested, consistent with an essential role for RNA unwinding in the repression of conjugation. The K37A/V38A, K39A/K40A, and Q41A/K42A mutants all exhibited defects in unwinding and duplexing comparable to W36A, yet none of the double mutants showed a significant loss of repressor activity *in vivo*. This observation may indicate that Trp-36 is also important for some other function,

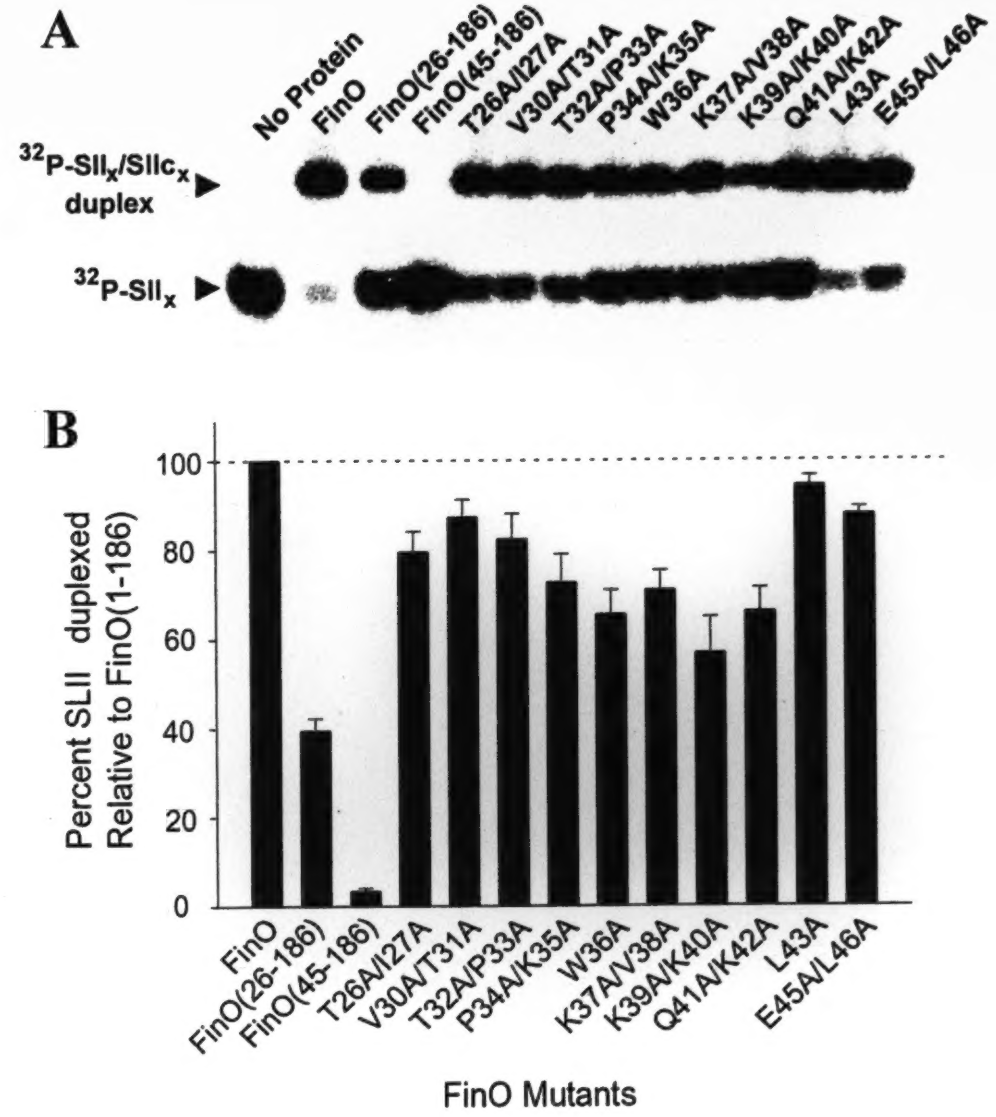


Figure 6.10: Duplexing efficiencies of N-terminal mutants. (**A**) Two hour duplexing reactions were performed in the presence of FinO and the various FinO mutants indicated. (**B**) The relative percentages of SLII duplexed are shown in graphical form.

other than RNA unwinding and sense-antisense RNA pairing, which ultimately contributes to the repression of plasmid transfer.

FinO has also been shown to stabilize FinP against endonucleolytic degradation by RNase E, a function that is also thought to contribute to the repression of conjugation (Frost *et al.*, 1989; Lee *et al.*, 1992; Jerome *et al.*, 1999). To test the effects of the N-terminal FinO mutations on FinP stability, we introduced the mutants into a FinP-expressing *E. coli* strain, and measured the stability of FinP transcripts isolated from these strains at various times after rifampicin-induced blockage of transcription (Figure 6.11). Consistent with previous results, FinP is rapidly degraded in cells that do not express FinO ($t1/2 < 5$ min), however, in cells that express wild type FinO, FinP is stable for at least 2 hours after cessation of transcription. Similar levels of stabilization are observed for all the N-terminal FinO mutants. We therefore conclude that the deficiencies of conjugation repression observed in these mutants is a direct consequence of their inability to unwind RNA and therefore facilitate FinP-*traJ* interactions, and cannot be explained by changes in FinP stabilization.

6.3 Discussion

In this study we have demonstrated that FinO has a novel RNA unwinding activity that utilizes its RNA binding free energy to destabilize base pairs. This unwinding activity appears to be required to denature otherwise stable secondary structures that exists both in its FinP and *traJ* mRNA targets, thereby facilitating sense-antisense base pairing between FinP and *traJ* RNA. This activity, together with the previously demonstrated ability of FinO to bind and protect FinP against endonucleolytic degradation, explains how this protein is able to enhance the FinP-mediated repression of conjugation 100-1000-fold *in vivo*.

The molecular mechanism underlying this activity is as yet unknown, however, an analysis of previous biochemical and structural data on this system, as well as data from other protein-RNA systems may provide insights into features of this mechanism. Our deletion study reveals that FinO residues 1-44 are absolutely required for RNA unwinding, but not for high affinity interactions with RNA. In contrast, further N-terminal or C-terminal deletions significantly decrease the affinity of FinO for RNA

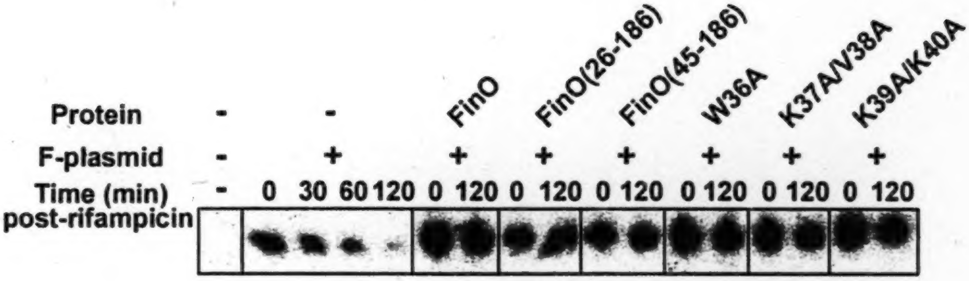


Figure 6.11: N-terminus of FinO is not required for FinP stabilization. Stabilization of FinP in cells expressing FinP and the indicated proteins was examined at the given times after the addition of rifampicin by northern blot analysis. As controls, FinP stability was examined in the absence of protein, and hybridization was performed on RNA extracted from cells not harboring the F-plasmid. This work was performed by Michael Gubbins.

(Chapter 3), indicating that the N-terminal region plays a particularly critical and direct role in RNA unwinding. The N-terminal ~32 residues of FinO appear to be disordered either in the free protein, or in complex with an RNA target, however, the C-terminal portion of this region (residues 33-44), becomes more structured upon binding RNA (Chapter 3) and direct interactions between this region and RNA have recently been demonstrated by site-specific protein-RNA crosslinking (see chapter 5). The structure of the free protein, crystallized at low temperature, revealed that residues 33-44 form the N-terminus of a lysine-rich, solvent exposed helix. Alanine point mutations within this region revealed that in addition to the lysines, several hydrophobic residues, most notably, Trp-36, as well as Pro-32 and Val-38, also mediate RNA unwinding. We propose that the N-terminal tip of this helix is docked to duplex RNA via electrostatic interactions between the lysine residues and the phosphodiester backbone of the RNA. This docking may position the critical hydrophobic residues in a groove of the RNA duplex, where they could potentially intercalate between adjacent bases and thereby disrupt base pairing. The stacking of aromatic amino acids with unpaired bases is a common feature of many protein-RNA interactions, and appears to play a role in the ATP-dependent disruption of base-pairing by helicases (Kim *et al.*, 1998; Velankar *et al.*, 1999; Marians, 2000).

We previously suggested that the N-terminal helix of FinO might interact with RNA in a manner similar to the binding of the N-terminal helix of the λN transcriptional anti-terminator to its box B RNA hairpin target (Legault *et al.*, 1998). The λN helix aligns in a roughly parallel manner to the box B hairpin such that Trp-18 stacks with an unpaired base at the apex of the GNRA-like loop. We proposed that FinO might also contact the loop region of an SLII RNA through stacking interactions involving Trp-36. Such an arrangement would allow the positively charged surfaces on the body of the protein to interact with the duplex and the single stranded tails at the opposite end of SLII. However, recent in-gel FRET measurements, which probed the relative proximities of specific regions of FinO and SII RNA, suggest that the tip of the N-terminal helix is in closer proximity to the 5' single stranded tail than to the opposite end of the duplex stem, where the loop occurs in SLII (see chapter 5). However, these measurements require the modification of the N-terminal helix with the fluorophore Texas Red, which may perturb the way in which the protein contacts RNA. Furthermore, these measurements do not rule out the possibility of transient, dynamic interactions between the FinO N-terminal helix and the loop region of the target RNA. Further, high resolution crystallographic studies currently underway in our lab will be required to understand the mechanistic details of how FinO binds and unwinds RNA.

While FinO-catalysed RNA unwinding is critical to the ability of FinO to facilitate sense-antisense RNA interactions, we believe that FinO must also bring the complementary RNA molecules into close proximity. Otherwise, the unwound RNAs would simply "snap back" to their intramolecularly base-paired states before contact could be made with the complementary RNA partner. Sense-antisense RNA interactions that regulate a variety of aspects of plasmid replication and transfer are thought to proceed via loop-loop or kissing complex intermediates enroute the final, paired state (Franch *et al.*, 1999). For example, the ColE1 protein Rom directly interacts with and stabilizes a kissing complex to facilitate RNA-RNA pairing (Eguchi *et al.*, 1990, Predki *et al.*, 1995) and, while Rom and FinO are not structurally related, it is tempting to speculate that FinO may carry out a similar function. Tethering FinP and *traJ* RNA together in this way would faciltate pairing of the two RNAs, once unwinding of the internal secondary structures had taken place. Interestingly, the destabilization of internal RNA secondary structure has recently been shown to play a key role in sense-antisense RNA recognition that mediates the expression of the DNA replication initiator protein, RepA, which controls replication of F-like plasmids (Kolb *et al.*, 2001). In this system, the RNAs involved, CopA and CopT, do not require a protein co-factor to interact, but contain several base pair mismatches in a pair of stem-loop structures that form an intermolecular kissing complex. These bulges are adjacent to the unpaired loops and have been shown to be necessary to allow the propagation of intermolecular base pairing from the initial kissing complex to a more fully paired and therefore stable structure containing a four-stranded junction. In comparison, SLII in FinP and it's compliment in *traJ* mRNA are almost completely base paired, and it is intriguing to speculate that FinO may specifically destabilize intramolecular base pairs near the loops to facilitate the expansion of intermolecular base pairing from an initial kissing complex. While we have demonstrated that FinO can facilitate RNA-RNA interactions between full length FinP and *traJ* mRNA, as well as between minimal SLII-SLIIc substrates, the interactions between the larger RNAs occur almost 100-fold faster than with the minimal RNAs. The full length RNAs also contain SLI and SLIc, which are shorter than SLII and SLIIc, and also contain a single base mismatch lacking in SLII/SLIIc. Thus, the less stable SLI/SLIc, as well as the additional single-stranded regions present in the larger RNAs, may play a critical role in anti-sense RNA-RNA pairing.

ATP-dependent DNA helicases often unwind large tracts of duplex DNA with significant energy input from ATP hydrolysis. In contrast, most RNA helicases only need to unwind short segments of duplex and therefore, processivity may not be absolutely required for many RNA remodeling processes. We have shown that FinO can unwind small duplex RNA structures and facilitate new RNA-RNA interactions without energy input from ATP hydrolysis. Perhaps the best candidate as a functional homolog of FinO is the nucleocapsid protein (NCp7) of HIV-1, which is involved in the maturation process of the retrovirus's dimeric genome. NCp7 binds to kissing stem-loop structures, facilitates their transition into an extended duplex in an ATP-independent manner and has been shown to possess strand annealing, as well as helix destabilizing activity (Tsuchihashi & Brown, 1994; Rein *et al.*, 1998; Williams *et al.*, 2001). Continuous melting and annealing of double stranded RNA, by NCp7 has been suggested to reduce the kinetic barriers in the transition of single to double stranded nucleic acids and to facilitate formation of the lowest energy product, which is the extended duplex.

Chapter 7

Summary and Future Perspectives

The extensive studies on FinO presented within this thesis have greatly enhanced our understanding of the molecular mechanism employed by FinO to inhibit bacterial conjugation.

We initially suggested that FinO functions similarly to Rom and enhances the rate of duplex formation by stabilizing kissing complexes (Chapter 3). Since FinO does not appear to dimerize, the stabilizing interactions would require one FinO to bind both FinP and traJ-mRNA, or would involve induced protein-protein interactions upon association of the complementary loops. The N-terminal region, particularly residues 1-26, was considered the prime candidate for mediating kissing complex stabilization. This was based on several lines of evidence. First, the N-terminus is dispensable for high affinity binding of FinO to individual stem-loops, but is required for efficient FinO mediated duplex rate enhancement (Chapters 3 and 4). Second, we speculated that the elongated structure of FinO allows the protein to span the entire length of the stem-loop structure (Chapter 3). Since the C-terminal region was shown by Sandercock et al. (1999) to protecting the linker of SLII from RNaseE degradation, we proposed a model where the C-terminus is near the base of the stem and residues 1-26 are at the opposite end near the loop, where they could stabilize the kissing complex. We also hypothesized that the basic regions on the core of FinO and around the tip of the N-terminal helix would both contact the negatively charged RNA, while the negative patch on the base of FinO would be oriented away from the RNA. This further restrained our model and we suggested that the front face of FinO (i.e. the side with the basic patch on the core) interacts with the stem-loop as shown in Figure 4.6B.

Although our initial model seemed reasonable, we had no direct evidence for positioning the N-terminus near the loop region. Subsequent crosslinking and gelFRET assays were designed in an attempt to better define the association of FinO with its target stem-loop and to specifically see if the N-terminus was located near the loop (Chapter 5). Crosslinking between specific residues on FinO and SLII confirmed that the basic regions of FinO contact the RNA and that the negative base of FinO is not near the RNA. GelFRET analyses showed that residues on the core of FinO, as well as in the C-terminal helix, were in close proximity to the base of SLII, but contrary to our model, the N-terminal helix also appeared to be in close contact with the stem base and not with the loop. However, the gelFRET assay only reveals the interactions occurring on a time-averaged scale and says nothing about the dynamics of a system. The N-terminus could make transient interactions with the loop region in a process requiring either a large conformational change to move the N-terminal tip (possibly through an induced fit mechanism) or through a shorter-lived binding event involving an additional FinO molecule. Our crosslinking results revealed that multiple FinO molecules could bind a single stem-loop structure and that the N-terminus was involved in formation of the higher order oligomeric complexes. In contrast to gelFRET, crosslinking might trap some of the transient interactions between the N-terminus and the loop.

At the same time that we were performing the gelFRET assays, we discovered that FinO could unwind double stranded RNA and that this unwinding activity was required for FinO promoted duplex formation and conjugative inhibition (Chapter 6). The unwinding activity required the presence of the N-terminal 45 residues of FinO. While the first N-terminal 26 residues are required for maximal unwinding efficiency, the most critical residues in the unwinding process were residues 32-42. There was a direct correlation between the ability of FinO, FinO(26-186) and FinO(45-186) to unwind SII and to promote duplex formation between $SLII_x$ and $SLIIc_x$ or FinP and *traJ*-mRNA. There was also a correlation between unwinding and conjugative inhibition, with FinO(45-186) having no ability to block conjugation. Although it now appears likely that FinO unwinding does not require the same loop-loop stabilization function observed with Rom, one would expect that FinO must hold the complementary strands together in some fashion to prevent stem-loops from snapping back together on themselves after being destabilized. Models for how FinO might promote duplex formation are shown in Figure 7.1 (FinO mediated unwinding occurring near the base of the stem) and Figure 7.2 (FinO mediated unwinding occurring in the stem near the loop).

Using site directed mutagenesis, we localized the essential unwinding function to residues 32-42 (Chapter 6). The presence of these residues greatly reduced the affinity of FinO for SLII, suggesting that FinO-RNA binding energy is directed to base-pair destabilization. It is possible that upon binding of the RNA substrate, the N-terminus is pushed between the base pairs of the stem. Tryptophan-36 was the single most critical residue for unwinding activity and mutating this residue to alanine increased binding

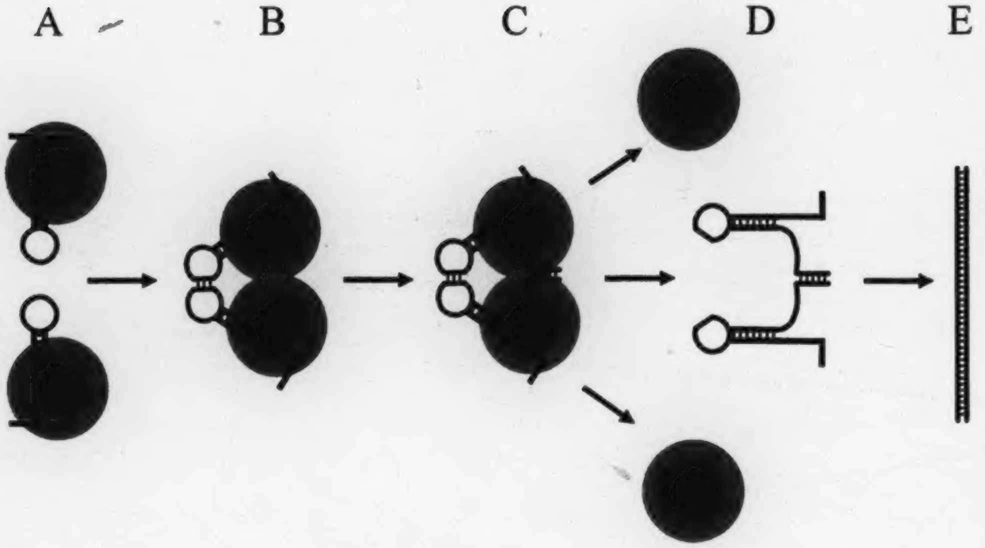


Figure 7.1: Model for how FinO might promote duplex formation between complementary stem-loop structures by unwinding the base of the stem. **(A)** FinO (shown in red) initially binds to the complementary stem-loop structures. **(B)** The complementary stem-loops first associate through base pairing at the loops to form a transient kissing intermediate. **(C)** FinO, functioning near the base of the stem-loop structures, can unwind RNA and can help to promote the stable association of the complementary single stranded regions in the two RNAs. **(D)** FinO dissociates from the RNA, the transient loop-loop interactions break apart and the stable duplex formed between the tails of the complementary RNAs is propagated through the rest of the RNAs to yield the final extended duplex shown in **(E)**.

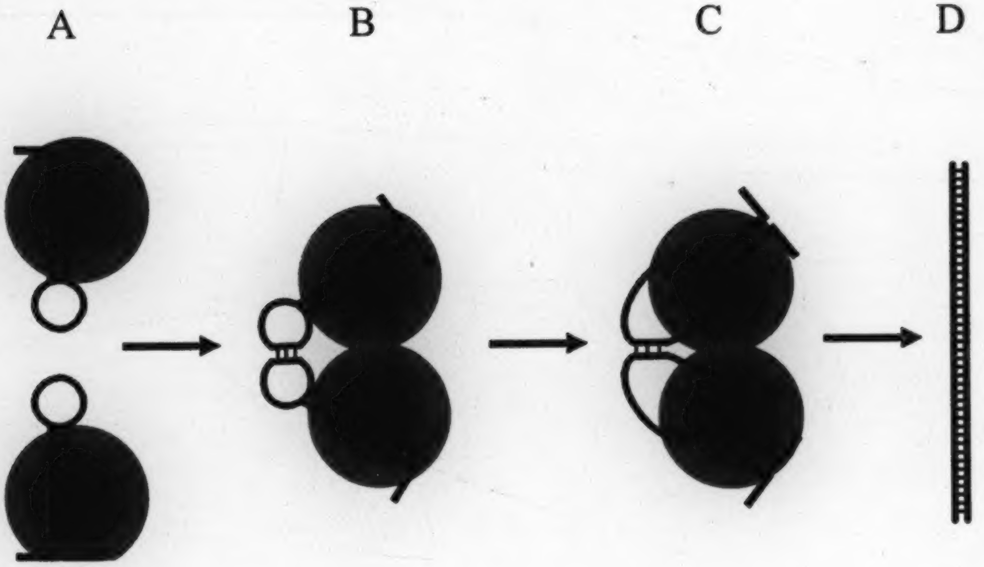


Figure 7.2: Model for how FinO might promote duplex formation between complementary stem-loop structures by unwinding the stem region below the loop. **(A)** FinO (shown in red) initially binds to the complementary stem-loop structures. **(B)** The complementary stem-loops first associate through base pairing at the loops to form a transient kissing intermediate. **(C)** FinO facilitates propagation of base pairing from the loops and through the stem region by unwinding the stem structure below the loops. **(D)** The duplex progresses through the entire sequence of the RNAs to yield the final extended duplex.

affinity of FinO for SLII by ~4 fold. We suggest that this residue along with other hydrophobic residues could destabilize base pairing by intercalating between the bases of the stem, thereby destroying the stacking arrangement required for stable base pairing. Several helicases are also thought to destabilize base-pairing interactions by insertion of hydrophobic residues between the stacked bases (Kim *et al.*, 1998; Velankar *et al.*, 1999; Marians, 2000). The use of a basic region for unfavorable interactions with RNA is quite intriguing given the fact that RNA binding proteins often use basic regions for enhancing binding to the negatively charged RNA. It is possible that many of the basic residues contribute to favorable binding with the negatively charged backbone of the RNA, while it is only the hydrophobic residues that promote the base-pair destabilization.

Interestingly, the region from residue 5 to 14 shares similar characteristics with the unwinding region of residues 32-42; there are 5 basic residues and 4 hydrophobic residues in each segment. These two regions may present a novel unwinding motif that may be a hallmark of other proteins with nucleic acid unwinding activity. Future studies should be directed towards determining if residues 5 to 14 are responsible for the unwinding activity observed in the first N-terminal 26 residues.

A potential candidate for a protein with functional homology to FinO is NCp7 from HIV. NCp7 possesses RNA chaperone activity, however, the mechanistic details underlying this activity are not well understood (Rein *et al.*, 1998). Although one role of NCp7 is to recognize stem-loop structures, it does not appear that binding of stem-loops is required for its chaperone activity (Takahashi *et al.*, 2001). The N-terminal 14 residues of NCp7 share some similarities with the unwinding region of FinO (both contain several basic residues and an aromatic amino acid) and although the binding requirements are different between FinO and NCp7, it will be interesting to see if this N-terminal region of NCp7 has any functional relationship to the basic N-terminal region of FinO.

In the absence or presence of FinO, the rate of duplexing between FinP and *traJ* mRNA is two orders of magnitude faster than duplexing between SLII and SLIIc (Chapter 6). SLI is therefore critical for obtaining maximal rates of unwinding. Due to their shorter stem length and the presence of a bulge in their stems 4 base pairs below the loop, SLI and SLIc is predicted to be much less stable, and therefore duplex much more rapidly, than SLII and SLIIc. Kolb *et al.* (2001) have recently shown that bulges in the stems of CopA and CopT help in promoting extended duplex formation between these RNA molecules. Binding of FinO to the high affinity SLII target occurs with an affinity ~ 10 fold higher than binding to SLI, therefore, FinO might specifically act to duplex SLII and SLIIc, while duplexing of the first stem loops might not depend on FinO but would still be required for obtaining the maximum overall duplexing rate between FinP and *traJ* mRNA. Duplexing assays between SLI and SLIc in the presence or absence of FinO would help better define the role of these structures in the overall duplexing process between FinP and *traJ* mRNA.

X-ray crystallographic studies are currently in progress to try and determine the three dimensional structure of FinO bound to stem-loop targets. The structural determination of this complex will enhance our understanding of how FinO recognizes its targets and where the unwinding event occurs. There are several other experiments in progress to try and determine where the base-pair destabilization occurs on SLII. One experiment involves using a duplex similar to SII, but with a covalent phosphate bond of one strand broken halfway up the stem. We will add FinO and see which half of this split strand is unwound first. Another experiment involves the use of NMR to map changes in imino proton shifts within SLII upon the addition of FinO. The imino protons of bases in the stem of SLII are involved in base-pairing interactions that protect the protons from exchange with deuterium and result in each imino protein having a characteristic proton shift. FinO base-pair destabilization would result in the disappearance of proton shifts from the bases that are released from base-pairing interactions.

The knowledge gained from studying the FinOP system as well as other antisense regulatory systems can be applied to the use of antisense gene therapies, where DNA or RNA oligos are introduced into cells to associate with target mRNAs and block translation (reviewed by Jen and Gewirtz, 2000). A problem with antisense gene therapies is the formation of secondary structures within the antisense strands. It would be interesting to see if FinO could function as a chaperone within eukaryotic cells to enhance the rate of duplexing between antisense genes and their targets.

Chapter 8

References

Achtman, M., Kennedy, N. & Skurray, R. (1977) Cell--cell interactions in conjugating *Escherichia coli*: role of TraT protein in surface exclusion *Proc Natl Acad Sci U S A* 74: 5104-5108.

Allers, J. & Shamoo, Y. (2001) Structure-based analysis of protein-RNA interactions using the program ENTANGLE *J Mol Biol* 311: 75-86.

Amarasinghe, G. K., De Guzman R. N. Turner R. B., Chancellor K.J., Wu Z. R. & Summers M. F. (2000) NMR structure of the HIV-1 nucleocapsid protein bound to stem-loop SL2 of the psi-RNA packaging signal. Implications for genome recognition *J Mol Biol* 301: 491-511.

Austin, S. & Nordstrom, K. (1990) Partition-mediated incompatibility of bacterial plasmids *Cell* 60: 351-354.

Bacterial conjugation, ed. Clewell, D. B. (1993) Plenum Press, Plenum New York.

Banner, D. W., Kokkinidis & M., Tsernoglou, D. (1987) Structure of the ColE1 Rop protein at 1.7 A resolution *J Mol Biol* 196: 657-675.

Battiste, J. L., Mao H., Rao N. S., Tan R., Muhandiram D. R., Kay L. E., Frankel A. D. & Williamson J. R. (1996) Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex *Science* 273: 1547-1551.

Bevilacqua, P. C. & Cech, T. R. (1996) Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR *Biochemistry* 35: 9983-9994.

Blomberg, P., Nordstrom, K. & Wagner, E. G. (1992) Replication control of plasmid R1: RepA synthesis is regulated by CopA RNA through inhibition of leader peptide translation *EMBO J* 11: 2675-2683.

Boyd, E. F. & Hartl, D. L. (1997) Recent horizontal transmission of plasmids between natural populations of *Escherichia coli* and *Salmonella enterica J Bacteriol* 179: 1622-1627.

Boyd, E. F. & Hartl, D. L. (1998) Salmonella virulence plasmid. Modular acquisition of the *spv* virulence region by an F-plasmid in *Salmonella enterica* subspecies I and insertion into the chromosome of subspecies II, IIIa, IV and VII isolates *Genetics* 149: 1183-1190.

Brunger, A. T. (1992) *X-PLOR: a system for X-ray crystallography and NMR* Yale University Press, New Haven, Connecticut.

Brünger, A. T. et al. (1999) Crystallography & NMR system: A new software suite for macromolecular structure determination *Acta Cryst.* D54: 905-921.

Burd, C. G. & Dreyfuss, G. (1994) Conserved structures and diversity of functions of RNA-binding proteins *Science* 265: 615-621.

Cai, Z., Gorin A., Frederick R., Ye X., Hu W., Majumdar A., & Kettani A (1998) Solution structure of P22 transcriptional antitermination N peptide- boxB RNA complex *Nat Struct Biol* 5: 203-212.

Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. (1999) X-ray crystal structures of 70S ribosome functional complexes *Science* 285: 2095-2104.

Cavarelli, J., Rees, B., Ruff, M., Thierry, J. C. & Moras, D. (1993) Yeast tRNA(Asp) recognition by its cognate class II aminoacyl-tRNA synthetase *Nature* 362: 181-184.

Cheah, K. C. & Skurray, R. (1986) The F plasmid carries an IS3 insertion within finO. *J. Gen. Microbiol.* 132: 3269-3275.

Chen, Y. & Ebright, R. H. (1993) Phenyl-azide-mediated photocrosslinking analysis of Cro-DNA interaction *J Mol Biol* 230: 453-460.

Clever, J. L., Wong, M. L. & Parslow, T. G. (1996) Requirements for kissing-loop-mediated dimerization of human immunodeficiency virus RNA *J Virol* 70: 5902-5908.

Comolli, L. R., Pelton, J. G. & Tinoco, I. Jr (1998) Mapping of a protein-RNA kissing hairpin interface: Rom and Tar-Tar* *Nucleic Acids Res* 26: 4688-4695.

Conn, G. L., Draper, D. E., Lattman, E. E. & Gittis, A. G. (1999) Crystal structure of a conserved ribosomal protein-RNA complex *Science* 284: 1171-1174.

Couturier, M., Bex, F., Bergquist, P. L. & Maas, W. K. (1988) Identification and classification of bacterial plasmids *Microbiol Rev* 52: 375-395.

Cowtan , K., *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* 31, 34-38 (94).

Cram, D. S., Loh, S. M., Cheah, K. C. & Skurray, R. A. (1991) Sequence and conservation of genes at the distal end of the transfer region on plasmids F and R6-5. *Gene* 104: 85-90.

Cuozzo M. & Silverman P.M. (1986) Characterization of the F plasmid TraJ protein synthesized in F' and Hfr strains of Escherichia coli K-12. *J. Biol. Chem.* 261: 5175-5179.

Datta, N. (1975) Epidemiology and classifiction of plasmids. *In* Schlessinger, D. (ed.) *Microbiology - 1974*. pp. 9-15. American Society for Microbiology, Washington, D.C.

Davis, B. D. (1950) Nonfiltrability of the agents of genetic recombination in *Escherichia coli*. *J Bacteriol* 60: 507-508.

De Guzman, R. N., Wu Z. R., Stalling C. C., Pappalardeo, L., Borer P. N. & Summers M. F. (1998) Structure of the HIV-1 nucleocapsid protein bound to the SL3 psi-RNA recognition element *Science* 279: 384-388.

Diebold, R. J., Rajaram, N., Leonard, D. A. & Kerppola, T. K. (1998) Molecular basis of cooperative DNA bending and oriented heterodimer binding in the NFAT1-Fos-Jun-ARRE2 complex *Proc Natl Acad Sci U S A* 95: 7915-7920.

Doublie, S. (1997) Preparation of selenomethionyl proteins for phase determination. *Meth. Enzymology* 276: 523-529.

Draper, D. E. (1999) Themes in RNA-protein recognition *J Mol Biol* 293: 255-270.

Eguchi, Y., Itoh, T. & Tomizawa, J. (1991) Antisense RNA *Annu Rev Biochem* 60: 631-652.

Eguchi, Y. & Tomizawa, J. (1990) Complex formed by complementary RNA stem-loops and its stabilization by a protein: function of ColE1 Rom protein *Cell* 60: 199-209.

Eisinger, J. & Spahr, P. F. (1973) Binding of complementary pentanucleotides to the anticodon loop of transfer RNA *J Mol Biol* 73: 131-137.

Esnouf, R. M. (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities *J Mol Graph Model* 15: 132-4, 112-113.

Faber, C., Scharpf, M., Becker, T., Sticht, H. & Rosch, P. (2001) The structure of the coliphage HK022 Nun protein-lambda-phage boxB RNA complex. Implications for the mechanism of transcription termination *J Biol Chem* 276: 32064-32070.

Feng, Y. X., Copeland, T. D., Henderson, L. E., Gorelick, R. J., Bosche, W. J., Levin, J. G & Rein, A. (1996) HIV-1 nucleocapsid protein induces "maturation" of dimeric retroviral RNA in vitro *Proc Natl Acad Sci U S A* 93: 7577-7581.

Feng, Y. X., Fu, W., Winter, A. J., Levin, J. G.& Rein, A. (1995) Multiple regions of Harvey sarcoma virus RNA can dimerize in vitro *J Virol* 69: 2486-2490.

Finlay, B. B., Frost, L. S. & Paranchych, W. (1986) Nucleotide sequences of the R1-19 plasmid transfer genes *traM*, *fin*P, *traJ*, and *traY* and the *traYZ* promoter *J Bacteriol* 166: 368-374.

Finnegan, D. J. & Willetts, N. S. (1972) The nature of the transfer inhibitor of several F-like plasmids. *Mol. Gen. Genet.* 119, 57-66.

Finnegan, D. J. and Willetts, N. S. (1971) Two classess of F*lac* mutants insensitive to transfer inhibition by an F-like R factor *Mol. Gen. Genet.* 111: 256-264.

Firth, N., Ippen-Ihler, K., & Skurray, R. A. (1996) Structure and function of the F factor and mechanism of conjugation *In Neidhardt, F. C. (ed.) Escherichia coli and Salmonella, (2nd Edition). American Society for Microbiology, Washingtion, D.C.*

Firth, N. & Skurray, R. (1992) Characterization of the F plasmid bifunctional conjugation gene, traG *Mol Gen Genet* 232: 145-153.

Franch, T., Petersen, M., Wagner, E. G. & Jacobsen, J. P. Gerdes, K. (1999) Antisense RNA regulation in prokaryotes: rapid RNA/RNA interaction facilitated by a general U-turn loop structure *J Mol Biol* 294: 1115-1125.

Frankel, A. D. & Young, J. A. (1998) HIV-1: fifteen proteins and an RNA *Annu Rev Biochem* 67: 1-25.

Freier, S. M. & Tinoco, I. Jr (1975) The binding of complementary oligoribonucleotides to yeast initiator Transfer RNA *Biochemistry* 14: 3310-3314.

Frost, L., Lee, S., Yanchar, N., & Paranchych, W. (1989) *finP* and *fisO* mutations in FinP anti-sense RNA suggest a model for FinOP action in the repression of bacterial conjugation by the Flac plasmid JCFL0. *Mol Gen Genet* 218, 152-160.

Frost, L. S., Ippen-Ihler, K., & Skurray, R. A., (1994) Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol. Rev.* 58, 162-210.

Gaudin, H. M. & Silverman, P. M. (1993) Contributions of promoter context and structure to regulated expression of the F plasmid traY promoter in Escherichia coli K-12 *Mol Microbiol* 8: 335-342.

Ghetu, A. F., Gubbins, M. J., Frost, L. S., & Glover, J. N. (2000) Crystal structure of the bacterial conjugation repressor FinO. *Nat Struct Biol* 7: 565-769.

Ghetu, A. F., Gubbins, M. J., Oikawa, K., Kay, C. M., Frost, L. S., & Glover, J. N. (1999) The FinO repressor of bacterial conjugation contains two RNA binding regions. *Biochemistry* 38: 14036-14044.

Gutell, R. R., Cannone, J. J., Konings, D. & Gautheret, D. (2000) Predicting U-turns in ribosomal RNA with comparative sequence analysis *J Mol Biol* 300: 791-803.

Hall, K. B. & Fox, R. O. (1999) Directed cleavage of RNA with protein-tethered EDTA-Fe. *Methods* 18, 78-84.

Hayes, W. (1953) Recombination in E. coli K12: Unidirectional transfer of genetic material *Nature* 169: 118-119.

Herschlag, D. (1995) RNA chaperones and the RNA folding problems *JBC* 270 (36): 20871-20874

Hjalt, T. & Wagner, E. G. (1992) The effect of loop size in antisense and target RNAs on the efficiency of antisense RNA control *Nucleic Acids Res* 20: 6723-6732.

Hjalt, T. A. & Wagner, E. G. (1995) Bulged-out nucleotides in an antisense RNA are required for rapid target RNA binding in vitro and inhibition in vivo *Nucleic Acids Res* 23: 580-587.

Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction *Gene* 77: 51-59

Holm, L. & Sander, C (1997) Protein folds and families: sequence and structure alignments *Nucleic Acids Res* 25, 231-234.

Houssier, C. & Grosjean, H. (1985) Temperature jump relaxation studies on the interactions between transfer RNAs with complementary anticodons. The effect of modified bases adjace587 t to the anticodon triplet *J Biomol Struct Dyn* 3: 387-408.

Iftode, C., Daniely, Y. & Borowiec, J. A. (1999) Replication protein A (RPA): the eukaryotic SSB *Crit Rev Biochem Mol Biol* 34: 141-180.

Ippen-Ihler, K. & Skurray, R. (1993) Genetic organization of transfer-related determinants on the sex factor F and related plasmids. In *Bacterial Conjugation* (Clewell, D., ed): 23-52, Plenum Press, New York.

Jankowsky, E., Gross, C. H., Shuman, S. & Pyle, A. M. (2000) The DExH protein NPH-II is a processive and directional motor for unwinding RNA *Nature* 403: 447-451.

Jen, K. Y. & Gewirtz, A. M. (2000) Suppression of gene expression by targeted disruption of messenger RNA: available options and current strategies *Stem Cells* 18: 307-319.

Jerome, L. J. (1999) PhD Thesis. University of Alberta. Edmonton, Alberta.

Jerome, L. J. & Frost, L. S. (1999) In vitro analysis of the interaction between the FinO protein and FinP antisense RNA of F-like conjugative plasmids *J Biol Chem* 274: 10356-10362.

Jerome, L. J., van Biesen, T. & Frost, L. S. (1999) Degradation of FinP antisense RNA from F-like plasmids: the RNA-binding protein, FinO, protects FinP from ribonuclease E *J. Mol. Biol.* 285: 1457-1473.

Jessen, T. H., Oubridge, C., Teo, C. H., Pritchard, C. & Nagai, K. (1991) Identification of molecular contacts between the U1 A small nuclear ribonucleoprotein and U1 RNA *EMBO J* 10: 3447-3456.

Ji, X., Klarmann, G. J. & Preston, B. D. (1996) Effect of human immunodeficiency virus type 1 (HIV-1) nucleocapsid protein on HIV-1 reverse transcriptase activity in vitro *Biochemistry* 35: 132-143.

Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard (1991) Improved methods for binding protein models in electron density maps and the location of errors in these models *Acta Crystallogr A* 47 (Pt 2): 110-119.

Jucker, F. M. & Pardi, A. (1995) GNRA tetraloops make a U-turn *RNA* 1: 219-222.

Karn, J. (1999) Tackling Tat *J Mol Biol* 293: 235-254.

Khan, R. & Giedroc, D. P. (1992) Recombinant human immunodeficiency virus type 1 nucleocapsid (NCp7) protein unwinds tRNA *J Biol Chem* 267: 6689-6695.

Kim, J. L., Morgenstern, K. A., Griffith, J. P., Dwyer, M. D., Thomson, J. A, Murcko, M. A., & Lin C, Caron (1998) Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding *Structure* 6: 89-100.

Knee, R. & Murphy, P. R. (1997) Regulation of gene expression by natural antisense RNA transcripts *Neurochem Int* 31: 379-392.

Kolb, F. A., Westhof, E., Ehresmann, C., Ehresmann, B., Wagner, E. G. & Romby, P. (2001) Bulged residues promote the progression of a loop-loop interaction to a stable and inhibitory antisense-target RNA complex *Nucleic Acids Res* 29: 3145-3153.

Koraimann, G., Koraimann, C., Koronakis, V., Schlager, S., & Hogenauer, G. (1991) Repression and derepression of conjugation of plasmid R1 by wild-type and mutated finP antisense RNA *Mol. Microbiol.* 5: 77-87.

Koraimann, G., Teferle, K., Markolin, G., Woger, W. & Hogenauer, G. (1996) The FinOP repressor system of plasmid R1: analysis of the antisense RNA control of traJ expression and conjugative DNA transfer *Mol. Microbiol.* 21: 811-821.

Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures *J. Applied Crystallogr.* 24: 946-950.

Lapadat-Tapolsky, M., Pernelle, C., Borie, C. & Darlix, J. L. (1995) Analysis of the nucleic acid annealing activities of nucleocapsid protein from HIV-1 *Nucleic Acids Res* 23: 2434-2441.

Laughrea, M., Jette, L., Mak, J., Kleiman, L., Liang, C. & Wainberg, M. A. (1997) Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization *J Virol* 71: 3397-3406.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein strctures *J. Appl. Crystallogr.* 26: 283-291

Lederberg, J., Cavalli, L. J. & Lederberg, E. M. (1952) Sex compatibility in *Escherichia coli Genetics* 37: 720-730.

Lederberg, J. & Tatum, E. (1946) Gene recombination in *E. coli Nature* 158: 558.

Lee, S. H., Frost, L. S. & Paranchych, W. (1992) FinOP repression of the F plasmid involves extension of the half-life of FinP antisense RNA by FinO *Mol. Gen. Genet.* 235: 131-139.

Legault, P., Li, J., Mogridge, J., Kay, L. E. & Greenblatt, J. (1998) NMR structure of the bacteriophage lambda N peptide/boxB RNA complex: recognition of a GNRA fold by an arginine-rich motif *Cell* 93: 289-299.

Leonard, D. A. & Kerppola, T. K. (1998) DNA bending determines Fos-Jun heterodimer orientation *Nat Struct Biol* 5: 877-881.

Lever, A., Gottlinger, H., Haseltine, W. & Sodroski, J. (1989) Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions *J Virol* 63: 4085-4087.

Lohman, T. M. & Ferrari, M. E. (1994) Escherichia coli single-stranded DNA-binding protein: multiple DNA- binding modes and cooperativities *Annu Rev Biochem* 63: 527-570.

Madej, T., Gibrat, J. F. & Bryant, S. H. (1995) Threading a database of protein cores *Proteins* 23: 356-369.

Marians, K. J. (2000) Crawling and wiggling on DNA: structural insights to the mechanism of DNA unwinding by helicases *Structure Fold Des* 8: R227-235.

Masai, H. & Arai, K. (1988) RepA protein- and *oriR*-dependent initiation of R1 plasmid replication: identification of a rho-dependent transcription terminator required for cis-action of repA protein *Nucleic Acids Res* 16: 6493-6514.

Masai, H., Kaziro, Y. & Arai, K. (1983) Definition of *oriR*, the minimum DNA segment essential for initiation of R1 plasmid replication in vitro *Proc Natl Acad Sci U S A* 80: 6814-6818.

Masukata, H. & Tomizawa, J. (1986) Control of primer formation for ColE1 plasmid replication: conformational change of the primer transcript *Cell* 44: 125-136.

Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999) Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure *J Mol Biol* 288: 911-940.

Matson, S. W., Bean, D. W. & George, J. W. (1994) DNA helicases: enzymes with essential roles in all aspects of DNA metabolism *Bioessays* 16: 13-22.

Mazodier, P. & Davies, J. (1991) Gene transfer between distantly related bacteria *Annu. Rev. Genet.* 25: 147-171.

McRee, D. E. (1999) XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density *J Struct Biol* 125: 156-65.

Merino, E., Balbas, P., Puente, J. L. & Bolivar, F. (1994) Antisense overlapping open reading frames in genes from bacteria to humans *Nucleic Acids Res* 22: 1903-1908.

Merritt, E. A. & Bacon, D. J. (1997) Raster3D: Photorealistic Molecular Graphics *Methods Enzymol.* 277: 505-524.

Miki, T., Horiuchi, T. & Willetts, N. S. (1978) Identification and characterization of four new *tra* cistrons on the E. coli K12 sex factor F *Plasmid* 1: 316-323.

Mullineaux, P. & Willetts, N. (1985) Promoters in the transfer region of plasmid F. *Basic Life Sci.* 30: 605-614.

Muriaux, D., De Rocquigny, H., Roques, B. P. & Paoletti, J. (1996) NCp7 activates HIV-1Lai RNA dimerization by converting a transient loop- loop complex into a stable dimer *J Biol Chem* 271: 33686-33692.

Nicholls, A., Sharp, K. A. & Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons *Proteins* 11: 281-96.

Novotny, C. P. & Fives-Taylor, P. (1974) Retraction of F pili *J Bacteriol* 117: 1306-11.

Ohki, M. & Tomizawa, J. (1968) Asymmetric transfer of DNA strands in bacterial conjugation *Cold Spring Harbor Symp Quant Biol* 33: 651-658.

Otwinowski, Z. & Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode *Meth. Enzymology* 276: 307-325.

Oubridge, C., Ito, N., Evans, P. R., Teo, C. H. & Nagai, K. (1994) Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin *Nature* 372: 432-438.

Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C. & Marquet, R. (1996) A loop-loop "kissing" complex is the essential part of the dimer linkage of genomic HIV-1 RNA *Proc Natl Acad Sci U S A* 93: 5572-5577.

Pendergrast, P. S., Chen, Y., Ebright, Y. W. & Ebright, R. H. (1992) Determination of the orientation of a DNA binding motif in a protein- DNA complex by photocrosslinking *Proc Natl Acad Sci U S A* 89: 10287-10291.

Penfold, S. S., Simon, J. & Frost, L. S. (1996) Regulation of the expression of the *traM* gene of the F sex factor of Escherichia coli *Mol Microbiol* 20: 549-558.

Persson, C., Wagner, E. G., Nordstrom, K. (1988) Control of replication of plasmid R1: kinetics of in vitro interaction between the antisense RNA, CopA, and its target, CopT *EMBO J* 7: 3279-3288.

Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994) Three-dimensional structure of a hammerhead ribozyme *Nature* 372: 68-74.

Pollard, V. W. & Malim, M. H. (1998) The HIV-1 Rev protein *Annu Rev Microbiol* 52: 491-532.

Pongs, O., Bald, R. & Reinwald, E. (1973) On the structure of yeast tRNA Phe. Complementary-oligonucleotide binding studies *Eur J Biochem* 32: 117-25.

Porter, D. J., Short, S. A., Hanlon, M. H., Preugschat, F., Wilson, J. E., Willard, D. H Jr. & Consler, T. G. (1998) Product release is the major contributor to kcat for the hepatitis C virus helicase-catalyzed strand separation of short duplex DNA *J Biol Chem* 273: 18906-18914.

Predki, P. F., Nayak, L. M., Gottlieb, M. B. & Regan, L. (1995) Dissecting RNA-protein interactions: RNA-RNA recognition by Rop *Cell* 80: 41-50.

Price, S. R., Evans, P. R. & Nagai, K. (1998) Crystal structure of the spliceosomal U2B"-U2A' protein complex bound to a fragment of U2 small nuclear RNA *Nature* 394: 645-650.

Provencher, S. W. & Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism*Biochemistry* 20: 33-37.

Puglisi, E. V. & Puglisi, J. D. (1998) HIV-1 A-rich RNA loop mimics the tRNA anticodon structure *Nat Struct Biol* 5: 1033-1036.

Puglisi, J. D., Chen, L., Blanchard, S. & Frankel, A. D. (1995) Solution structure of a bovine immunodeficiency virus Tat-TAR peptide- RNA complex *Science* 270: 1200-1203.

Ramirez-Carrozzi, V. & Kerppola, T. (2001) Gel-based fluorescence resonance energy transfer (gelFRET) analysis of nucleoprotein complex architecture *Methods* in press:

Ramirez-Carrozzi, V. R. & Kerppola, T. K. (2001) Long-range electrostatic interactions influence the orientation of Fos- Jun binding at AP-1 sites *J Mol Biol* 305: 411-427.

Ramirez-Carrozzi, V. R. & Kerppola, T. K. (2001) Dynamics of Fos-Jun-NFAT1 complexes *Proc Natl Acad Sci U S A* 98: 4893-4898.

Ramirez-Carrozzi, V. R. & Kerppola, T. K. (2001) Control of the orientation of fos-jun binding and the transcriptional cooperativity of Fos-Jun-NFAT1 complexes *J Biol Chem*

276: 21797-21808.

Rein, A., Henderson, L. E. & Levin, J. G. (1998) Nucleic-acid-chaperone activity of retroviral nucleocapsid proteins: significance for viral replication *Trends Biochem Sci* 23: 297-301.

Rupp, W. D. & Ihler, G. (1968) Strand selection during bacterial mating *Cold Spring Harbor Symp Quant Biol* 33: 647-650.

Ryter, J. M. & Schultz, S. C. (1998) Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA *EMBO J* 17: 7505-7513.

Sandercock, J. R. and Frost, L. S. (1998) Analysis of the major domains of the F fertility inhibition protein, *Mol. Gen. Genet.* 259: 622-629.

Sarathy, P. V. & Siddiqi, O. (1973) DNA synthesis during bacterial conjugation. II. Is DNA replication in the Hfr obligatory for chromosome transfer? *J Mol Biol* 78: 443-451.

Scharpf, M., Sticht, H., Schweimer, K., Boehm, M., Hoffmann, S. & Rosch, P. (2000) Antitermination in bacteriophage lambda. The structure of the N36 peptide-boxB RNA complex *Eur J Biochem* 267: 2397-2408.

Schwab, M., Reisenzein, H. & Hogenauer, G. (1993) TraM of plasmid R1 regulates its own expression *Mol Microbiol* 7: 795-803.

Shahied, L., Braswell, E. H., LeStourgeon, W. M. & Krezel, A. M. (2001) An antiparallel four-helix bundle orients the high-affinity RNA binding sites in hnRNP C: a mechanism for RNA chaperonin activity *J Mol Biol* 305: 817-828.

Silver, L. L. & Bostian, K. A. (1993) Discovery and development of new antibiotics: the problem of antibiotic resistance *Antimicrob. Agents Chemother.* 37: 377-383.

Soultanas, P., Dillingham, M. S., Wiley, P., Webb, M. R. & Wigley, D. B. (2000) Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism *EMBO J* 19: 3799-3810.

Staley, J. P. & Guthrie, C. (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things *Cell* 92: 315-326.

Stallings, S. C. & Moore, P. B. (1997) The structure of an essential splicing element: stem loop IIa from yeast U2 snRNA *Structure* 5: 1173-1185.

Stolt, P. & Zillig, W. (1993) Antisense RNA mediates transcriptional processing in an archaebacterium, indicating a novel kind of RNase activity *Mol Microbiol* 7: 875-882.

Su, L., Radek, J. T., Hallenga, K., Hermanto, P., Chan, G., Labeots, L. A. & Weiss, M. A. (1997) RNA recognition by a bent alpha-helix regulates transcriptional antitermination in phage lambda *Biochemistry* 36: 12722-12732.

Takahashi, K., Baba, S., Koyanagi, Y., Yamamoto, N., Takaku, H. & Kawai, G. (2001) Two basic regions of NCp7 are sufficient for conformational conversion of HIV-1 dimerization initiation site from kissing-loop dimer to extended-duplex dimer *J Biol Chem* 276: 31274-31278.

Tanner, N. K. & Linder, P. (2001) DExD/H box RNA helicases: from generic motors to specific dissociation functions *Mol Cell* 8: 251-262.

Terwilliger, T. C., Kim, S.-H., Eisenberg, D., (1987) Generalized method of determining heavy-atom positions using the difference Patterson function *Acta Crystallogr. A Foundations of Crystallography.*43: 1-5.

Thompson, R., Taylor, L., Kelly, K., Everett, R. & Willetts, N. (1984) The F plasmid origin of transfer: DNA sequence of wild-type and mutant origins and location of origin-specific nicks *EMBO J* 3: 1175-1180.

Tomizawa, J. (1990) Control of ColE1 plasmid replication. Intermediates in the binding of RNA I and RNA II *J Mol Biol* 212: 683-694.

Tomizawa, J. & Itoh, T. (1981) Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript *Proc Natl Acad Sci U S A* 78: 6096-6100.

Tomizawa, J., Itoh, T., Selzer, G. & Som, T. (1981) Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA *Proc Natl Acad Sci U S A* 78: 1421-1425.

Tomizawa, J. & Som, T. (1984) Control of ColE1 plasmid replication: enhancement of binding of RNA I to the primer transcript by the Rom protein *Cell* 38: 871-878.

Tsuchihashi, Z., Khosla, M. &Herschlag, D. (1993) Protein enhancement of hammerhead ribozyme catalysis *Science* 262: 99-102.

Turner, B. G. & Summers, M. F. (1999) Structural biology of HIV *J Mol Biol* 285: 1-32.

van Biesen, T. & Frost, L. S. (1992) Differential levels of fertility inhibition among F-like plasmids are related to the cellular concentration of *finO* mRNA. *Mol. Microbiol.* 6: 771-780.

van Biesen, T. & Frost, L. S. (1994) The FinO protein of IncF plasmids binds FinP antisense RNA and its target, traJ mRNA, and promotes duplex formation *Mol. Microbiol.* 14: 427-436.

van Biesen, T., Soderbom, F., Wagner, E. G. & Frost, L. S. (1993) Structural and functional analyses of the FinP antisense RNA regulatory system of the F conjugative plasmid. *Mol. Microbiol.* 10: 35-43.

Vapnek, D. & Rupp, W. D. (1970) Asymmetric segregation of the complementary sex-factor DNA strands during conjugation in *Escherichia coli J Mol Biol* 53: 287-303.

Velankar, S. S., Soultanas, P., Dillingham, M. S., Subramanya, H. S. & Wigley, D. B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism *Cell* 97: 75-84.

von Hippel, P. H. & Delagoutte, E. (2001) A general model for nucleic acid helicases and their "coupling" within macromolecular machines *Cell* 104: 177-190.

Wagner, E. G. & Simons, R. W. (1994) Antisense RNA control in bacteria, phages, and plasmids *Annu Rev Microbiol* 48: 713-742.

Wagner, E. G., von Heijne, J. & Nordstrom, K. (1987) Control of replication of plasmid R1: translation of the 7k reading frame in the RepA mRNA leader region counteracts the interaction between CopA RNA and CopT RNA *EMBO J* 6: 515-522.

Watanabe, T. & Fukasawa, T. (1961) Episome-mediated transfer of drug resistance in Enterobacteriaeea. I. Transfer of resistance factors by conjugation *J Bacteriol* 81: 669-678.

Westhof, E., Dumas, P. & Moras, D. (1988) Restrained refinement of two crystalline forms of yeast aspartic acid and phenylalanine transfer RNA crystals *Acta Crystallogr A* 44 (Pt 2): 112-23.

Wilkins, B. & Lanka, E. (1993) DNA processing and replication during plasmid transfer between gram negative bacteria. In *Bacterial Conjugation* (Clewell, D., ed): 105-136, Plenum Press, New York.

Willetts, N. (1972) Location of the origin of transfer of the sex factor F *J Bacteriol* 112: 773-778

Willetts, N. & Maule, J. (1974) Interactions between the surface exclusion systems of some F-like plasmids *Genet Res* 24: 81-89.

Willetts, N. & Maule, J. (1986) Specificities of IncF plasmid conjugation genes *Genet. Res.* 47: 1-11.

Williams, M. C., Rouzina, I.., Wenner, J. R., Gorelick, R. J., Musier-Forsyth, K. & Bloomfield, B. A. (2001) Mechanism for nucleic acid chaperone activity of HIV-1 nucleocapsid protein revealed by single molecule stretching *Proc Natl Acad Sci U S A* 98: 6121-6126.

Womble, D. D., Dong, X., Wu, R. P., Luckow, V. A., Martinez, A. F. & Rownd, R. H.. (1984) IncFII plasmid incompatibility product and its target are both RNA transcripts *J Bacteriol* 160: 28-35.

Woody, R. W. (1995) Circular dichroism. *Methods Enzymol.* 246: 34-71.

Wu, W., Henderson, L. E., Copeland, T. D., Gorelick, R. J., Bosche, W. J., Rein, A. & Levin, J. G. (1996) Human immunodeficiency virus type 1 nucleocapsid protein reduces reverse transcriptase pausing at a secondary structure near the murine leukemia virus polypurine tract *J Virol* 70: 7132-7142.

Yoon, K., Turner, D. H. & Tinoco, I. Jr (1975) The kinetics of codon-anticodon interaction in yeast phenylalanine transfer RNA *J Mol Biol* 99: 507-518.

Yoshioka, Y., Fujita, Y. & Ohtsubo, E. (1990) Nucleotide sequence of the promoter-distal region of the tra operon of plasmid R100, including *traI* (DNA helicase I) and *traD* genes. *J. Mol. Biol.* 214: 39-53.

Yoshioka, Y., Ohtsubo, H. & Ohtsubo, E. (1987) Repressor gene *finO* in plasmids R100 and F: constitutive transfer of plasmid F is caused by insertion of IS3 into F *finO*. *J. Bacteriol.* 169: 619-623.